A meta-analytic review of the Motivated Strategies for Learning Questionnaire

L. Alison Phillips, Rutgers University
Marcus Crede, University at Albany, State University of New York

Available at: https://works.bepress.com/lalison-phillips/9/
A meta-analytic review of the Motivated Strategies for Learning Questionnaire

Marcus Credé a,⁎, L. Alison Phillips b

a Department of Psychology, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY 12222, United States
b Department of Psychology, Rutgers, The State University of New Jersey, 112 Paterson Street, New Brunswick, NJ 08901, United States

Article history:
Received 21 July 2010
Received in revised form 25 January 2011
Accepted 12 March 2011
Available online xxxx

Abstract

The current paper presents a meta-analytic review of the Motivated Strategies for Learning Questionnaire (MSLQ), which consists of fifteen subscales designed from classic social-cognitive learning theories and which is widely used to predict academic performance. Results based on 2158 correlations from 67 independent samples and 19,900 college students indicate that the subscales of the MSLQ vary in their utility for predicting grades, with grade-related validities ranging from $\rho = 0.40$ for the subscale measuring students’ effort regulation to $\rho = 0.05$ for the subscale measuring students’ help-seeking behaviors. Factor-analyses of the meta-analytic intercorrelations broadly support the theoretical structure of the MSLQ. Alteration or elimination of items with undesirable psychometric characteristics could potentially both augment empirical support for the theoretical structure of the MSLQ and strengthen its subscales’ predictive utility for academic performance.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Researchers have predominantly focused on stable student traits, abilities, and behaviors to predict academic performance in college. For example, a substantial literature (e.g., Halpin, Halpin, & Schaer, 1981; Hezlett et al., 2001; Kuncel, Hezlett, & Ones, 2001) has provided evidence of the relatively strong prediction of college grades by scores on standardized admissions tests, which theoretically assess a student’s stable cognitive ability, which reflects the ability to process, understand, and remember new information (Humphreys, 1979). Others (e.g., Lievens, Coetsier, De Fruyt, & De Maeseneer, 2002; O’Connor & Faunnonen, 2007; Poropat, 2009; Robbins et al., 2004) have examined the relationship between academic performance and theoretically stable traits, such as personality (e.g., conscientiousness) and motivational factors (e.g., need for achievement), that are thought to directly affect important academic behaviors such as study skills and class attendance, which are themselves strongly predictive of college grades (e.g., Credé & Kuncel, 2008; Credé, Roehm, & Kiesczynka, 2010). While traits, abilities, and skills explain a substantial proportion of the variation in academic performance, research in these areas have been less successful at providing insights into the specific manner in which students acquire new knowledge. One approach that has focused specifically on the manner in which students engage with academic tasks and material and has sought to shed further light onto both the determinants of academic performance and the processes of learning is represented by the self-regulated learning literature (e.g., Zimmerman, 1990; Zimmerman & Schunk, 1989; Rotgans & Schmidt, 2009).

Students who engage in self-regulated learning are defined as “...metacognitively, motivationally, and behaviorally active participants in their learning” (Zimmerman, 1990, p. 4). That is, students who engage in self-regulated learning monitor their own progress toward self-set goals and are therefore able to reflect on the effectiveness of their learning approaches (i.e., have high levels of meta-cognition), tend to view the learning task as intrinsically interesting and worthwhile while having high levels of self-efficacy (i.e., have learning appropriate cognitions and motivations), and engage in and persist with learning behaviors that maximize the degree to which learning occurs (i.e., engage in learning appropriate behaviors). All three components of self-regulated learning (meta-cognitions, motivations, and behaviors) are assumed to be important determinants of learning and hence academic performance, although the effect of meta-cognition and motivations on academic performance are typically presented as being mediated through learning behaviors (e.g., Duncan & McKeachie, 2005), such that meta-cognition and appropriate motivations result in the use of appropriate learning strategies that, in turn, have a positive effect on academic performance.

An important theoretical addition to the theory of self-regulated learning came from researchers and theorists (e.g., Duncan & McKeachie, 2005; Dweck & Leggett, 1988; Pintrich, 2000; Schunk, 2005a, 2005b; Zimmerman, 1989) who advocated a social–cognitive view of learning, which suggests that self-regulated learning exhibits...
not only between-person variation but also within-person variation. That is, self-regulation is seen as situation specific, and self-regulated learning can therefore be used to explain why one student performs better on an academic task than another student (i.e., explain between-person variation in academic performance), but can also be used to explain why a single student performs better on one academic task than on other academic tasks (i.e., explain within-person variation). While meta-cognitive and self-regulatory abilities are likely to be relatively constant across both time and academic tasks due to their trait-like nature, social–cognitive theories of learning suggest that both the motivations and cognitions relating to a task (e.g., judgments regarding the intrinsic value of a task), and the learning behaviors used by a student to perform that task can vary substantially across tasks. For example, a student’s motivations, cognitions, and learning behaviors for academic tasks may vary across different classes (e.g., advanced seminar in chosen major versus a required general education class), and even across different tasks within the same class (e.g., studying for a multiple choice exam versus writing a term paper). This within-person variation in learning processes suggests need for a contextualized (task specific) examination of self-regulated learning processes.

Empirical investigations of self-regulated learning processes and mechanisms require the existence of valid and reliable measures of the various constructs that are central to self-regulated learning. This need led Paul Pintrich and colleagues to develop the Motivated Strategies for Learning Questionnaire (MSLQ, Pintrich, Smith, Garcia & McKeachie, 1991) — a single measure designed to assess task-specific cognitions and motivations as well as the learning strategies used by students to engage with that task. Tens of thousands of students have since been evaluated on the MSLQ; a search of the PsychINFO database reveals over 150 articles that have used the MSLQ and the MSLQ manual has itself been cited over 700 times. The MSLQ has been used to study college students in a wide variety of countries (e.g., Suksaamran, 2003; Thongnoum, 2002), settings (e.g., community colleges, Welsh, 2007; online classes, Gaythwaite, 2006), and populations (e.g., teaching students, Bhattacharyya, 2004; nursing students, Tutor, 2006). However, the use of the MSLQ in these studies is almost always as a static predictor of academic performance: the ‘context-specificity of the constructs included in the MSLQ as implied by social–cognitive views of learning have not yet been rigorously tested except for a single recent paper (Rotgans & Schmidt, 2009) that showed significant variation in self-regulated learning processes across classes. Reviewing the results of studies that have used the MSLQ would, however, provide valuable information on the validity of the more general theory of self-regulated learning and provide information as to whether the MSLQ is likely to be a valuable tool in future examinations of the context-specificity of learning processes. Furthermore, testing the validity of the assumption that scores on the MSLQ normatively predict academic performance (at a static level) is important for educational practitioners and college counselors who may use the MSLQ for the early identification of at-risk students.

To this aim, the current paper has three primary purposes. The first is to meta-analyze the many studies that have investigated the predictive utility of the MSLQ subscales for academic performance. Findings regarding the validity of the MSLQ scales vary widely, with some authors finding impressive relationships between academic achievement and scores on the MSLQ (e.g., Bell, 2006; Bembenutty, 2007; Hsu, 1997; Langley, 2007) and others finding much weaker relationships (e.g., Barker, 1997; Bartels & Magun-Jackson, 2009; Huang, 2008; Lewis, 2006). The meta-analysis will yield a more comprehensive and extensive picture of the subscales’ validity for normative, ‘static’ use and for testing more complex theoretical relationships regarding the changeability and context-specificity of students’ motivations and learning strategies. The second purpose of the current paper is to find preliminary evidence for the notion that the MSLQ subscales tap context-specific student traits and behaviors by comparing the utility of the MSLQ subscales for predicting course-specific performance to their utility for predicting general academic performance (GPA). The third purpose is to highlight potentially problematic psychometric properties of some of the MSLQ items that may be limiting the subscales’ predictive utility and ability to contribute to learning theories. The MSLQ has been used in non-college settings (e.g., Pintrich & DeGroot, 1990) but its most frequent use has been in a college setting and we therefore focus specifically on the relationship between scores on MSLQ subscales and academic performance in college.

2. Overview of the MSLQ

The MSLQ is an 81-item self-report measure comprised of fifteen subscales that assess both student motivation to engage with course material and their learning strategies. The full MSLQ takes 20–30 min to complete, although individual subscales can also be administered (Pintrich, Smith, Garcia, & McKeachie, 1991). All items are Likert scale items and students use a seven-point response option format (1 = “Not at all true of me”; 7 = “Very true of me”) to respond to each item. Eight of the 81-items are reverse scored.

There are a total of thirty-one item and six subscales that attempt to assess students’ motivational orientations toward a college course. These six scales were designed to capture three theoretical components of motivation, termed ‘value beliefs’, ‘expectancy’, and ‘affect’ (Duncan & McKeachie, 2005). Value beliefs are assessed with three motivation subscales: a four-item intrinsic goal orientation scale that assesses whether a student’s participation in the academic task is an end in itself and whether the student is focused on the mastery of material, a four-item extrinsic goal orientation scale that assesses the degree to which a student participates in the task in order to obtain good grades, rewards or approval, and a six-item task value scale that assesses the degree to which the student has a belief that the class material is interesting and worth learning. Expectancy is assessed with two subscales: self-efficacy (confidence that a task can be performed) and control beliefs (akin to an internal locus of control); both capture the degree to which students believe that they have control over their level of achievement in a class. Finally, the affective component of motivation is measured by the test anxiety subscale, which attempts to capture the degree to which students experience anxiety and fear when taking tests.

The learning strategies subscales are comprised of fifty items divided into nine subscales that were designed to capture three theoretical types of learning strategies (Duncan & McKeachie, 2005). First, cognitive strategies are assessed with four subscales that measure rehearsal (repeating information to oneself), elaboration (summarizing information), organization (organizing new information using tables and outlines), and critical thinking (evaluating new ideas and applying them to novel situations). Metacognitive strategies are assessed by a single subscale that attempts to capture students’ ability to monitor their own mental processes and adjust them when needed — an important component of self-regulatory frameworks of learning (e.g., Pintrich, 2000). Lastly, resource management is assessed with four subscales that attempt to capture students’ ability to manage the resources available to them. These include time and study environment regulation (study skills and appropriate use of study time), effort regulation (persisting in studying even when material is boring), peer learning (working with others and using study groups), and help seeking (asking for assistance from peers or instructors). An overview of the structure of the MSLQ, along with example items is provided in Table 1. Readers interested in a more detailed account of the development of the MSLQ or the items for each of these subscales are referred to Pintrich et al. (1991), Pintrich, Smith, Garcia, and McKeachie (1993), and Duncan and McKeachie (2005).
3. Methodology

3.1. Search strategy

We identified potential data sources via keyword searches of the PsychINFO, Eric, and Dissertation Abstracts databases as well as internet searches and examinations of the reference lists of identified articles. Dissertations judged to be likely to include relevant data were purchased from Proquest. We also attempted to contact authors of articles that did not present information in a manner that allowed inclusion in this review (criteria described below) in order to obtain the relevant data. Each data source was coded by one of the two authors using standardized coding sheets that cued coders to attend to specific study characteristics that were identified as important for meta-analytic purposes. This information included: correlations between scores on MSLQ subscales and academic performance, correlations among MSLQ subscales, reliability information for scores on MSLQ subscales, the design of the study (predictive versus concurrent), the measure of academic performance (e.g., grades in a single class versus GPA), the type of class in which the information was gathered (e.g., Chemistry, English, and Mathematics), the source of academic performance ratings (self-ratings or instructor records), and a variety of other sample characteristics (e.g., gender, ethnicity, age, and language of administration). This information was subsequently entered in a database to facilitate the appropriate sorting of information by the relationship captured by the correlation, criterion category, and type of research design.

The coding of studies for meta-analysis has previously been shown to exhibit high inter-rater reliability especially when coding decisions are easy to make (Whetzel & McDaniel, 1988; Zakzanis, 1998), but a random sample of these articles was nevertheless coded by both authors to allow for an examination of coding accuracy. Agreement in the coding of relevant study characteristics was almost unanimous; the few disagreements resulted from articles providing discrepant data between tables and text (e.g., sample size in correlation table not corresponding with sample size given in description of methodology). All such disagreements were resolved via discussion.

3.2. Inclusion and exclusion criteria

Studies were included if they were described in English and reported zero-order correlations between MSLQ subscales and either GPA or grades in an individual class for college students, or reported data in a manner that allowed such correlations to be calculated (e.g., MSLQ scores between high and low performing students). Studies were excluded if they reported non-Pearson correlations or if they selectively reported only statistically significant correlations (i.e., did not report any of their non-significant results), as the inclusion of such studies would have yielded biased meta-analytic estimates of the relationships among variables. Studies were also excluded if results were based on dramatically shortened or reworded versions of the MSLQ or if the results reported in the studies suggested flaws in the manner in which the MSLQ was scored (e.g., reporting negative alpha reliabilities for some MSLQ subscales, a likely artifact of items that had not been reverse scored). We also screened our final database carefully to ensure that data that might have been published in two separate venues (e.g., once as a dissertation and then as a journal article) was not included twice in our analysis.

For studies that utilized predictive designs with more than one assessment of MSLQ subscales (e.g., MSLQ subscales administered at time 1 and time 2 with academic performance gathered at time 2), we coded both the predictive and concurrent correlations (for the moderator analysis; see Section 3.6) and took the average of the two correlations for the overall analysis.

3.3. Criterion categories

Only two broad criteria were considered in this meta-analysis. The MSLQ was specifically designed to assess motivation and learning strategies in a particular class and we therefore examined the relationship between scores on MSLQ subscales and performance in an individual class as our primary analysis. The MSLQ is founded on the assumption (Duncan & McKeachie, 2005) that motivational variables and learning strategies are changeable across tasks (e.g., Illies & Judge, 2005), an assumption shown to be valid by Rotgans and Schmidt (2009) who found significant within-person variation in
MSLQ scores across three subjects (English, Math, and Science). At the same time these classes also exhibit some stability across situations (e.g., classes) for the same individual (e.g., Bong, 2001, 2004; Gottfried, 1985; Vermeeten, Lodewijks, & Vermunt, 1997; Warr & Downing, 2000), and so a number of authors have either used the MSLQ to measure general tendencies to use certain learning strategies (e.g., Wolters, 2003) or have presented correlations between MSLQ scores and performance across multiple classes (i.e., GPA). We therefore also present analyses for the relationship between scores on MSLQ subscales and college GPA. Comparing predictive utility of the MSLQ subscales for GPA versus course-specific grades allows for initial investigation into the theoretical claim that motivation and learning strategies may be context-specific.

3.4. Final databases

In total, the database consisted of 2158 correlations from 67 independent samples (from a total of 59 articles) representing 19,900 independent college students. The database for the relationships between MSLQ subscales and GPA consisted of 98 correlations from 24 independent samples representing 9696 independent college students. The database for the relationships between MSLQ subscales and grades in individual classes consisted of 431 correlations from 53 independent samples representing 15,321 independent college students. The database for the intercorrelations among MSLQ subscales consisted of 1629 correlations from 35 independent samples representing 11,507 independent college students.

3.5. Analytic strategy

We used the Hunter and Schmidt (2004) interactive random-effects meta-analytic method to aggregate the findings of the examined studies. Meta-analytic methods that combine study-level effects are preferable to a simple pooling of data since this avoids the potential for introducing Simpson’s paradox (i.e., reversing the direction of effects observed at a subgroup level when aggregating data across the subgroups) into research findings. The Hunter and Schmidt method is preferable to alternative meta-analytic methods (e.g., Rosenthal & Rubin, 1982, method, or Hedges, 1982, method), for two primary reasons. First, it provides an estimate of the relationship that would be observed between variables in the absence of study artifacts such as unreliability in measurement of the dependent and independent variables, range restriction, or artificial dichotomization. Second, it provides information as to the presence of possible unexamined moderators of the relationship by providing an estimate of the variability in observed effect sizes after accounting for the variability that is due to sampling error and other study artifacts (e.g., differential reliability of measurement).

The interactive meta-analytic method is appropriate when only some studies report artifact information. Specifically, it involves using the reported artifact information to construct artifact distributions that are then used to correct the distribution of observed effect sizes for measurement artifacts. For this meta-analysis we constructed artifact distributions for unreliability in the measurement of each of the MSLQ subscales and for unreliability in the measurement of the academic performance criterion. Given that very few studies in the educational literature have ever reported reliability information for GPA and none (to our knowledge) have reported reliability information for individual grades, we imported the four published reliability estimates known to us for GPA (Barritt, 1966; Bendig, 1953; Reilly & Warech, 1993; Stricker, Rock, Burton, Muraki, & Jirele, 1994), while estimates of grade reliability were taken from five large undergraduate classes recently taught by the first author. Summary information for the various reliability distributions is provided in Table 2. Mean reliabilities for scores on the MSLQ subscales range from a low of $\alpha = .59$ (help seeking) to a high of $\alpha = .91$ (self-efficacy).

The Hunter and Schmidt method provides three primary pieces of information: $\rho_{rob}$, the mean sample-size weighted observed correlation; $\rho$, the sample-size weighted estimate of the true score correlation after correcting for study artifacts; and $SD_{\rho_{rob}}$, the standard deviation of true score correlations. $SD_{\rho_{rob}}$ is, in turn, used as evidence for the presence of moderators of the relationship, as much as it represents the amount of variability across studies that remains after accounting for study artifacts. Large $SD_{\rho_{rob}}$ values are interpreted as evidence for the likely presence of moderators. $SD_{\rho_{rob}}$ values are also used for the construction of credibility intervals around the estimate of the true score correlation. Credibility intervals represent the range of plausible values for the true score correlation across situations. Wide credibility intervals suggest the presence of moderators. The Hunter and Schmidt (2004) software (version 1.1) used for this analysis also provides an estimate of the proportion of the observed variance that can be accounted for by examined study artifacts. Hunter and Schmidt (2004) argue that moderators are unlikely to exist if the proportion of variance that is accounted for by examined artifacts exceeds 75% since most meta-analyses cannot account for all likely study artifacts. This meta-analysis, for example, was not able to account for likely range restriction artifacts. Finally, we also reported operational validities that represent an estimate of the relationship that would be observed between the scores on a scale and a theoretical criterion measured with perfect reliability. That is, we corrected correlations for unreliability in the criterion but not for unreliability in the predictor since educational practitioners would have to use the MSLQ as it currently stands and would not be able to assess the MSLQ constructs with perfect reliability.

Lastly, the meta-analytic intercorrelation matrix of MSLQ subscales was used to examine the structure of the MSLQ in more detail. Specifically, the full meta-analytic inter-correlation matrix was factor analyzed (principal axis factoring and oblique rotation) using exploratory factor analysis (EFA) based on a harmonic mean sample size of 4230.

3.6. Moderators

Only two possible moderators, of the relationship between MSLQ subscales and academic performance were examined; both integral to a better understanding of self-regulated learning processes. First, we examined the difference between studies that used predictive designs (i.e., MSLQ subscales gathered at time 1 and academic performance at

<table>
<thead>
<tr>
<th>MSLQ subscales</th>
<th>$M_{r_{rob}}$</th>
<th>$SD_{r_{rob}}$</th>
<th>$k_{rob}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic Goal Orientation</td>
<td>.69</td>
<td>.05</td>
<td>21</td>
</tr>
<tr>
<td>Extrinsic goal orientation</td>
<td>.66</td>
<td>.10</td>
<td>16</td>
</tr>
<tr>
<td>Task value</td>
<td>.87</td>
<td>.04</td>
<td>17</td>
</tr>
<tr>
<td>Control of learning beliefs</td>
<td>.65</td>
<td>.11</td>
<td>14</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>.91</td>
<td>.02</td>
<td>21</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>.77</td>
<td>.06</td>
<td>12</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>.68</td>
<td>.05</td>
<td>14</td>
</tr>
<tr>
<td>Elaboration</td>
<td>.76</td>
<td>.05</td>
<td>19</td>
</tr>
<tr>
<td>Organization</td>
<td>.70</td>
<td>.07</td>
<td>14</td>
</tr>
<tr>
<td>Critical thinking</td>
<td>.77</td>
<td>.04</td>
<td>16</td>
</tr>
<tr>
<td>Metacognitive self-regulation</td>
<td>.77</td>
<td>.06</td>
<td>29</td>
</tr>
<tr>
<td>Time and study environment</td>
<td>.72</td>
<td>.07</td>
<td>15</td>
</tr>
<tr>
<td>Effort regulation</td>
<td>.61</td>
<td>.10</td>
<td>14</td>
</tr>
<tr>
<td>Peer learning</td>
<td>.68</td>
<td>.10</td>
<td>12</td>
</tr>
<tr>
<td>Help seeking</td>
<td>.59</td>
<td>.12</td>
<td>13</td>
</tr>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>.80</td>
<td>.07</td>
<td>4</td>
</tr>
<tr>
<td>Grades</td>
<td>.74</td>
<td>.08</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: MSLQ = Motivated Strategies for Learning Questionnaire; $M_{r_{rob}} = \text{mean of reliability distribution}$; $SD_{r_{rob}} = \text{standard deviation of reliability distribution}$; and $k_{rob} = \text{number of independent reliability coefficients on which distributions are based}$. 

Table 2: Reliability artifact distributions for MSLQ subscales, college GPA, and grades.
time 2) versus concurrent designs. Given that students are likely to base their responses to MSLQ subscales in part on their known academic performance, we expected correlations to be higher for concurrent designs than for predictive designs. That is, a student who has done poorly in a class is more likely to attribute his/her poor performance to factors such as poor strategies or low motivation rather than low capability (Schunk, 2005a) – particularly when that student has good self-regulation skills. The resultant higher concordance between achieved grades and self-rated motivation and strategies will increase correlations in concurrent designs above those observed in predictive designs. Second, we examined the difference in correlations for studies that used GPA as the measure of academic performance and studies that used the grade in a specific class as the measure of academic performance. Since the MSLQ is phrased to assess motivation and strategies for a specific class and since a student’s motivation and strategies are likely to be somewhat different for different classes, we expected correlations to be higher for studies that used the grade in an individual class as a measure of academic performance.

3.7. Critique of meta-analysis

Criticisms of meta-analytic methods have typically focused on the subjective nature of the process whereby studies are selected for inclusion in the meta-analysis, the possibility that meta-analytic findings quickly become outdated as new studies are published (e.g., Gay, Mills, & Airasian, 2009), and the possibility of drawing inappropriate inferences due to Simpson’s Paradox and the Ecological Fallacy (e.g., Cooper & Patall, 2005). Although these issues do pose potential threats to the validity of meta-analytic findings, the methods employed in this study have attempted to minimize these threats. Subjectivity in study selection was largely avoided by including all available sources of data, even going so far as to purchase numerous dissertations in order to minimize the file-drawer effect, and only screening out studies with obvious data-analytic errors (e.g., reporting of negative alpha reliabilities). The Ecological Fallacy would occur if we (or authors of the summarized articles) had inappropriately used group level characteristics to draw inferences about individuals. However, all of the examined data sources were based on students’ individual grades and students’ individual scores on the MSLQ and the Ecological Fallacy therefore does not represent a likely threat. Simpson’s Paradox is perhaps more likely since it is plausible that grade–MSLQ relationships are substantially moderated by unexamined group variables, but the Hunter and Schmidt meta-analytic method precludes Simpson’s paradox from being introduced since it involves the computation of sample-size weighted average correlations rather than the pooling of data across samples followed by the computation of a single correlation. It is however possible that Simpson’s paradox is present at the level of the individual studies included in our review. That is, unexamined group moderators of MSLQ–grade relationships may exist such that an examination of data at the overall sample level suggests relationships in opposite direction to those that would be found if the data was examined at the appropriate sub-group level. A number of authors (e.g., Bembenutty, 2007) had examined the effect of group moderators such as gender and ethnicity on the validity of MSLQ scores (i.e., MSLQ–grade relationships) but found no evidence for strong effects. These examinations, together with the relatively high level of congruence between our findings and the theoretically expected effects suggest that our findings are relatively unlikely to be characterized by Simpson’s paradox.

4. Results

4.1. Sample characteristics

Data was drawn from a wide variety of college populations. Samples were drawn from 7 countries: the United States, Canada, Australia, Thailand, Taiwan, India, and Malaysia (87.3% of the studies were conducted in North America) and a wide variety of class types: economics, science, mathematics, medicine and English courses. A sample-size weighted review of available study characteristics indicated that 37.1% of participants were male, 67.6% were Caucasian, 17.8% were Asian, 10.5% were African-American, and 6.1% were Hispanic. The sample-size weighted mean age participants was 23.1 years.

4.2. Criterion correlations

Meta-analytic results for the relationships between MSLQ subscales and both GPA and grades in individual classes are presented in Table 3. As hypothesized, the relationships between scores on MSLQ subscales and grades in individual classes were higher than those observed between scores on MSLQ subscales and GPA for all fifteen subscales (the relationship between peer learning and GPA was not in the expected direction). The relationships between MSLQ scores and academic performance were generally weak to moderate (Cohen, 1988). The highest observed validities for grades in individual classes were found for effort regulation ($\rho = .40, k=24, N=5180$), self-efficacy ($\rho = .37, k=39, N=8123$), and time and study environment ($\rho = .31, k=24, N=4892$). Similarly, the highest observed validities for GPA were found for effort regulation ($\rho = .23, k=5, N=2721$), and meta-cognitive self-regulation ($\rho = .22, k=13, N=4390$). Especially notable were the findings that specific learning approaches (e.g., rehearsal, elaboration, organization, and peer learning) were largely unrelated to academic performance while the theoretically more stable (i.e., less contextual) abilities and tendencies such as meta-cognitive self-regulation and effort regulation were most strongly related to academic performance.

The size of SDp estimates (and hence the width of the credibility intervals) was somewhat variable but large enough to suggest the possible presence of moderators for at least some MSLQ subscales (e.g., test anxiety and effort regulation). The type of class may influence the degree to which constructs assessed by the MSLQ influence academic performance in that class. For example, classes in which grades are based less on tests taken under exam conditions and more on take-home assignments and essays may not exhibit a strong relationship between test anxiety and academic performance.

4.3. Prospective and concurrent designs

Results for the moderation analysis examining the differences in relationships between scores on MSLQ subscales and grades in individual classes for studies using prospective versus concurrent designs are presented in Table 4. Contrary to expectations, the correlations for concurrent study designs were not uniformly higher than those for prospective study designs; concurrent designs exhibited higher correlations for only six of the fifteen subscales. Interestingly, all six of these subscales were motivational subscales with the largest observed difference observed for test anxiety ($\rho = -.42$ for concurrent designs versus $\rho = -.14$ for prospective designs), self-efficacy ($\rho = .58$ versus $\rho = .31$) and task value ($\rho = .42$ versus $\rho = .19$).

4.4. Intercorrelations

Meta-analytic estimates of the relationships among the fifteen subscales are presented in Table 5. Most disattenuated correlations (i.e., $\rho$) are of a medium to large size (Cohen, 1988), and the pattern of intercorrelations was generally supportive of the assumption that cognitions and motivations influence learning strategies. That is, cognitions and motivations exhibited moderate to large correlations with the various learning strategies. It is important to note that some of the meta-analytic intercorrelations suggested non-trivial construct

overlapping. Specifically, disattenuated correlations close to one were observed for the relationships between time and study environment and effort regulation ($\rho = .95$, $k = 16$, $N = 4833$), and peer learning and help seeking ($\rho = .92$, $k = 14$, $N = 4583$), while a number of other disattenuated correlations were also high enough to suggest some non-trivial level of construct redundancy (e.g., the relationship between elaboration and organization; $\rho = .82$, $k = 23$, $N = 6533$).

The full matrix of disattenuated correlations was used to conduct an EFA of the structure of MSLQ scales. Parallel analysis (Hayton, Allen, & Scarpello, 2004) for the EFA suggested a four-factor solution, and an examination of factor loadings (see Table 6) suggests that the first and second factors are composed of the learning strategies, the third factor is composed of motivational factors, and the fourth factor largely reflects test anxiety. This structure is largely supportive of the theoretical distinction between motivations and cognitions (factors 3 and 4) on the one hand and learning strategies (factors 1 and 2) on the other hand.

4.5. Review of items

Given the relatively low observed validities for many of the individual MSLQ subscales, we also conducted a review of the individual items that comprise the MSLQ. Specifically, we asked two independent raters with PhD-level training in psychometrics to evaluate the eighty-one individual items listed in Duncan and McKeeachie (2005) for two undesirable item characteristics: first, conditional-content items, and second, ideal-point items. Items with conditional content assume that a particular event has taken place and present a second statement with which the respondent must agree or disagree (e.g., “Whenever X occurs, I do Y”). Such items are psychometrically problematic because disagreement with such items could imply that X has never occurred for the respondent or that the respondent does not do Y when X occurs. Ideal-point items are items where both high-achieving and low-achieving students respond to the item in one way while average students are likely to respond to the item in a different way. The item, “I ask the instructor to clarify concepts I don’t understand well”, would probably elicit a disagreement from high performing students who do not require help, as well as from low performing students who might not be motivated enough to ask for help or lack the ability to recognize that they are struggling; average students who encounter occasional academic difficulties but recognize the need to ask for help may agree with the statement. Ideal-point items can be desirable when attempting to measure intermediate levels of a construct with high fidelity but require different scoring methods (Drasgow, Chernyshenko, & Stark, 2010), and do not lend themselves to the type of analyses typically conducted with MSLQ scores (e.g., factor analyses and correlations with grade outcomes).

Of the eighty-one examined items both raters independently identified ten items characterized by conditional content and eight items with ideal-point characteristics. These potentially problematic items disproportionately represent the peer learning and help seeking constructs. All three of the peer-learning items were classified as either having conditional item content (once) or having likely ideal-point characteristics (twice). Similarly, three out of the four items making up the help-seeking scale were classified as both conditional item content and being likely ideal-point items. These possible psychometric issues may help to explain some of the low validities reported in this meta-analytic review.
relationships with college academic performance. Moderate to strong relationships (Cohen, 1988) were observed between class grades and self-efficacy, effort regulation, and time and study environment subscales, although most other relationships between grades and MSLQ constructs were weaker. Encouragingly, the strongest of these relationships (e.g., $r = .41$ for effort regulation) are similar to those observed in meta-analytic reviews of some of the best known traditional predictors of academic performance: scores on admission tests, prior academic performance, and study skills and study habits (Credé & Kuncel, 2008; Hezlett et al., 2001).

More importantly, our results are broadly supportive of some of the basic assumptions that underpin theories of self-regulated learning (e.g., Zimmerman, 1990). That is, students who are able to engage in metacognitive self-monitoring and effort regulation, who view the academic task as having intrinsic interest and value, have high levels of self-efficacy, and who use appropriate learning strategies are shown to have higher average grades than students without these attributes and behaviors. Further, the pattern of relationships among the constructs assessed by the MSLQ supports the notion that the effect of motivations on academic performance is mediated by learning strategies in the manner suggested by the authors of the MSLQ (e.g., Duncan & McKeachie, 2005). That is, motivational variables assessed by the MSLQ are related to learning strategies, many of which are, in turn, related to academic performance.

### Table 4
Meta-analytic comparison of MSLQ validities from prospective and concurrent research designs.

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Study design</th>
<th>N</th>
<th>$k$</th>
<th>$r_{obs}$</th>
<th>SD$_{obs}$</th>
<th>$\rho$</th>
<th>SD$_{\rho}$</th>
<th>Lower 90%</th>
<th>Upper 90%</th>
<th>% Var explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic goal orientation</td>
<td>Prospective</td>
<td>8722</td>
<td>31</td>
<td>.13</td>
<td>.06</td>
<td>.33</td>
<td>.43</td>
<td>.62</td>
<td>.79</td>
<td>.64</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1086</td>
<td>5</td>
<td>.28</td>
<td>.13</td>
<td>.39</td>
<td>.18</td>
<td>.17</td>
<td>.62</td>
<td>.21</td>
</tr>
<tr>
<td>Extrinsic goal orientation</td>
<td>Prospective</td>
<td>8219</td>
<td>27</td>
<td>.06</td>
<td>.12</td>
<td>.08</td>
<td>.16</td>
<td>.13</td>
<td>.29</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>885</td>
<td>4</td>
<td>.13</td>
<td>.05</td>
<td>.18</td>
<td>.07</td>
<td>.09</td>
<td>.27</td>
<td>.64</td>
</tr>
<tr>
<td>Task value</td>
<td>Prospective</td>
<td>6289</td>
<td>28</td>
<td>.15</td>
<td>.09</td>
<td>.19</td>
<td>.11</td>
<td>.05</td>
<td>.34</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>885</td>
<td>4</td>
<td>.33</td>
<td>.11</td>
<td>.42</td>
<td>.13</td>
<td>.25</td>
<td>.58</td>
<td>.26</td>
</tr>
<tr>
<td>Control of learning beliefs</td>
<td>Prospective</td>
<td>4748</td>
<td>24</td>
<td>.12</td>
<td>.05</td>
<td>.42</td>
<td>.07</td>
<td>.08</td>
<td>.26</td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1115</td>
<td>5</td>
<td>.17</td>
<td>.00</td>
<td>.24</td>
<td>.00</td>
<td>.24</td>
<td>.24</td>
<td>.100</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>Prospective</td>
<td>6166</td>
<td>29</td>
<td>.26</td>
<td>.15</td>
<td>.31</td>
<td>.18</td>
<td>.09</td>
<td>.54</td>
<td>.17</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1172</td>
<td>6</td>
<td>.47</td>
<td>.00</td>
<td>.58</td>
<td>.00</td>
<td>.58</td>
<td>.58</td>
<td>.100</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>Prospective</td>
<td>5793</td>
<td>24</td>
<td>.11</td>
<td>.13</td>
<td>.42</td>
<td>.00</td>
<td>.42</td>
<td>.42</td>
<td>.100</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>716</td>
<td>3</td>
<td>.32</td>
<td>.01</td>
<td>.42</td>
<td>.00</td>
<td>.42</td>
<td>.42</td>
<td>.100</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>Prospective</td>
<td>4123</td>
<td>19</td>
<td>.10</td>
<td>.09</td>
<td>.14</td>
<td>.13</td>
<td>.03</td>
<td>.30</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1462</td>
<td>3</td>
<td>.06</td>
<td>.06</td>
<td>.08</td>
<td>.08</td>
<td>.02</td>
<td>.19</td>
<td>.38</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Prospective</td>
<td>5137</td>
<td>22</td>
<td>.13</td>
<td>.12</td>
<td>.17</td>
<td>.15</td>
<td>.03</td>
<td>.37</td>
<td>.24</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1516</td>
<td>4</td>
<td>.10</td>
<td>.15</td>
<td>.14</td>
<td>.19</td>
<td>.11</td>
<td>.38</td>
<td>.11</td>
</tr>
<tr>
<td>Organization</td>
<td>Prospective</td>
<td>3654</td>
<td>18</td>
<td>.07</td>
<td>.08</td>
<td>.10</td>
<td>.11</td>
<td>.04</td>
<td>.23</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1516</td>
<td>4</td>
<td>.04</td>
<td>.02</td>
<td>.06</td>
<td>.03</td>
<td>.02</td>
<td>.09</td>
<td>.88</td>
</tr>
<tr>
<td>Critical thinking</td>
<td>Prospective</td>
<td>3551</td>
<td>19</td>
<td>.10</td>
<td>.14</td>
<td>.13</td>
<td>.18</td>
<td>.10</td>
<td>.36</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1115</td>
<td>4</td>
<td>.06</td>
<td>.12</td>
<td>.08</td>
<td>.16</td>
<td>.13</td>
<td>.29</td>
<td>.19</td>
</tr>
<tr>
<td>Meta-cognitive self-regulation</td>
<td>Prospective</td>
<td>5223</td>
<td>23</td>
<td>.18</td>
<td>.12</td>
<td>.24</td>
<td>.16</td>
<td>.04</td>
<td>.44</td>
<td>.23</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>2041</td>
<td>8</td>
<td>.09</td>
<td>.16</td>
<td>.12</td>
<td>.20</td>
<td>.14</td>
<td>.38</td>
<td>.14</td>
</tr>
<tr>
<td>Time and study environment</td>
<td>Prospective</td>
<td>2919</td>
<td>15</td>
<td>.23</td>
<td>.06</td>
<td>.31</td>
<td>.07</td>
<td>.22</td>
<td>.41</td>
<td>.62</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1526</td>
<td>6</td>
<td>.22</td>
<td>.18</td>
<td>.30</td>
<td>.24</td>
<td>.00</td>
<td>.61</td>
<td>.11</td>
</tr>
<tr>
<td>Effort regulation</td>
<td>Prospective</td>
<td>3112</td>
<td>16</td>
<td>.28</td>
<td>.07</td>
<td>.41</td>
<td>.10</td>
<td>.29</td>
<td>.54</td>
<td>.54</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1719</td>
<td>6</td>
<td>.27</td>
<td>.22</td>
<td>.41</td>
<td>.32</td>
<td>.80</td>
<td>.80</td>
<td>.79</td>
</tr>
<tr>
<td>Peer learning</td>
<td>Prospective</td>
<td>2760</td>
<td>15</td>
<td>.11</td>
<td>.09</td>
<td>.09</td>
<td>.15</td>
<td>.10</td>
<td>.28</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1004</td>
<td>2</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.100</td>
</tr>
<tr>
<td>Help seeking</td>
<td>Prospective</td>
<td>2822</td>
<td>15</td>
<td>.05</td>
<td>.04</td>
<td>.08</td>
<td>.06</td>
<td>0</td>
<td>.16</td>
<td>.77</td>
</tr>
<tr>
<td></td>
<td>Concurrent</td>
<td>1227</td>
<td>4</td>
<td>.04</td>
<td>.00</td>
<td>.05</td>
<td>.00</td>
<td>.05</td>
<td>.05</td>
<td>.100</td>
</tr>
</tbody>
</table>

Note. MSLQ = Motivated Strategies for Learning Questionnaire; $N$ = number of participants; $k$ = number of studies; $r_{obs}$ = sample size weighted mean observed correlation; SD$_{obs}$ = observed standard deviation; $\rho$ = true score correlation; SD$_{\rho}$ = standard deviation of true score correlation; lower 90% = Lower bound of 90% credibility interval based on operational validity; upper 90% = upper bound of 90% credibility interval based on operational validity; % Var explained = the percentage of variance in the observed correlation accounted for by measurement error.
Our factor analytic findings are also broadly supportive of the theory of self-regulated learning (Zimmerman, 1990; Pintrich, 1999) inasmuch as they show a clear distinction between motivational and affective processes on one hand (factor 3 and factor 4, respectively) and learning approaches (factor 1 and factor 2) on the other hand. Our findings are also supportive of social–cognitive learning theories inasmuch as they support the contextual nature of self-regulated learning processes (e.g., Schunk, 2005b). That is, that motivations and learning strategies are class-specific and likely to vary across classes for the same individual; scores on the MSLQ were consistently more strongly related to performance in a specific class than to students’ average performance across multiple classes. We hope that these results will encourage researchers to further pursue investigations into the determinants of within-person variation in academic performance. We are aware of only one study in the MSLQ literature (Rotgans & Schmidt, 2009) that has formally examined the issue of within-person variability in MSLQ scores and the ability of this variation to account for within-person variation in grades. This relative dearth of research on the determinants of within-person variation in learning and academic achievement is, of course, not limited to the MSLQ literature but is a more general feature of research on academic performance in college settings.

One set of findings of our meta-analytic review that is less supportive of the general validity of the MSLQ and (to a lesser degree) the theory of self-regulated learning is that many of the specific learning strategies (i.e., rehearsal, elaboration, organization, critical thinking, peer learning, and help seeking) appeared to be largely unrelated to academic performance. Two possible reasons for these low relationships present themselves. First, we suspect that some of the learning strategies, particularly peer learning and help-seeking may exhibit curvilinear relationships with grades that are not captured by correlation coefficients. That is, neither very high performing nor very low performing students are likely to engage in peer learning or help seeking—high performing students because they have no need for help and low performing students because they do not know where to find help or are insufficiently motivated to do so. Rehearsal is another strategy that may not be related to academic performance in a linear manner. Rehearsal, as measured by the MSLQ, reflects rote memorization (2 items) and repetition of material (2 items)—two behaviors that are unlikely to be practiced by either high achieving students, who likely have no need for it, or low achieving students, who may be too unmotivated to engage in these behaviors or do not realize that they need to engage in these behaviors. The second possible reason for the low relationships between some learning strategies and academic performance is that the manner in which grades are determined in college often does not require students to use effective learning strategies. For example, performance on multiple choice examinations may require memorization but not require much critical thinking and students who engage in critical thinking may therefore not necessarily score more highly on tests. This would, in turn, result in trivial correlations between academic performance in the class and critical thinking behaviors. Future research on the predictive utility of scores on the MSLQ subscales with different types of performance tasks is warranted.

A review of the items that comprise the MSLQ also suggests that low correlations between the scales and academic performance may be due to poorly constructed items. Specifically, we identified a number of items with what we have termed conditional item content. That is, the item assumes that an event has occurred, presents the respondent with a response to the events, and asks the respondent to indicate whether he/she has engaged in this response. For example, one item asks respondents to indicate a level of endorsement for the statement: “When I can’t understand the material in this course, I ask another student in this class for help”. This item attempts to assess help-seeking behavior, but it is not clear if a disagreement with the item indicates that the respondent would not seek help or has never not understood the material.

It is also important to note that the MSLQ appears to be characterized by a non-trivial amount of redundancy at the measurement level. Scores on at least two pairs of subscales (time and study environment and effort regulation; peer learning and help seeking) are so strongly related with each other ($\rho = .92$ and $\rho = .95$, respectively) as to suggest that the same construct is being assessed. A rewriting of items to better distinguish between these pairs of constructs or a combination of these subscales may well be warranted.

Despite the identified psychometric problems and possible construct redundancy of the MSLQ, the main findings of our review suggest that the MSLQ is likely to be a valuable tool in both research and practice settings. The MSLQ appears to capture many of the most important constructs that are central to self-regulated learning and should therefore be valuable for future investigations of self-regulated learning in general and the contextual nature of self-regulated learning processes in particular (e.g., Rotgans & Schmidt, 2009). At a practical level, the MSLQ is likely to be useful for counselors attempting to understand the learning approaches of individual students and thereby target interventions to potentially problematic approaches (e.g., an overreliance on rehearsal strategies). Some researchers (e.g., Wolters, 2003) have even used the MSLQ to identify general approaches to learning and academic tasks (as opposed to the specific case of a single class), suggesting that the MSLQ could be used by researchers and instructors to make students more aware of their general approaches to learning and use it to engage students in a discussion of the efficacy of different learning strategies.

6. Limitations and future research

The MSLQ is explicitly based upon the notion that students are differently motivated and use different learning strategies for different courses and different academic tasks (Duncan & McKeachie, 2005), but we are not aware of any explicit examination of the validity of this assumption beyond the recent paper by Rutgans and Schmidt (2009). Capturing such within-person variation in motivations, behaviors, and learning strategies should substantially increase our ability to explain variance in academic performance above and beyond the variation explained by more stable student characteristics and behaviors, and also help researchers to better understand the act of learning (Schunk, 2005a). We therefore encourage future research to focus on the degree of variability in motivations and learning strategies (i.e., MSLQ scores) across academic tasks and courses. Similarly, we encourage further investigations into whether specific course characteristics may moderate the relationships of MSLQ constructs with academic performance. Rehearsal and elaboration strategies may, for example, be more important in classes that rely on multiple-choice tests for student assessment while critical thinking and help-seeking behaviors may be more important for courses that are difficult or theoretical in content.

Table 6

<table>
<thead>
<tr>
<th>MSLQ scales</th>
<th>Factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
</tr>
<tr>
<td>1 Intrinsic goal orientation</td>
<td>0.07</td>
</tr>
<tr>
<td>2 Extrinsic goal orientation</td>
<td>0.08</td>
</tr>
<tr>
<td>3 Task value</td>
<td>0.15</td>
</tr>
<tr>
<td>4 Control of learning beliefs</td>
<td>0.08</td>
</tr>
<tr>
<td>5 Self-efficacy</td>
<td>0.05</td>
</tr>
<tr>
<td>6 Test anxiety</td>
<td>0.04</td>
</tr>
<tr>
<td>7 Rehearsal</td>
<td>0.63</td>
</tr>
<tr>
<td>8 Elaboration</td>
<td>0.52</td>
</tr>
<tr>
<td>9 Organization</td>
<td>0.67</td>
</tr>
<tr>
<td>10 Critical thinking</td>
<td>0.18</td>
</tr>
<tr>
<td>11 Metacognitive self-regulation</td>
<td>0.61</td>
</tr>
<tr>
<td>12 Time and study environment</td>
<td>0.79</td>
</tr>
<tr>
<td>13 Effort regulation</td>
<td>0.65</td>
</tr>
<tr>
<td>14 Peer learning</td>
<td>−0.11</td>
</tr>
<tr>
<td>15 Help seeking</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Note. Loadings are from pattern matrix based on EPA with principal axis factoring and oblimin rotation.
The observed variation in effect sizes are certainly large enough to suggest the presence of some unexamined moderator variables.

We also encourage further research on how and when the motivations and learning strategies assessed by the MSLQ are developed. Some encouraging work has already been done that illustrates that learning strategies and motivations can change in response to direct interventions (e.g., work on the Learning to Learn courses, Van der Stoep & Pintrich, 2003), but we encourage further examinations of their stability, whether they are already fixed by the time that students arrive in college, and whether or not they change in response to performance on individual test or assignments. We also encourage researchers to examine the relationship of constructs assessed by the MSLQ with other widely used predictors of academic performance such as scores on entrance examinations, study behaviors, and stable student attributes such as personality. Some (e.g., Bidjenero & Dai, 2007) have already begun such examinations, but significantly more work in this area is required to establish the unique contribution of the constructs assessed by the MSLQ over more stable predictors of academic performance. Further, we hope that some of the possible non-linear relationships between learning strategies and academic performance that we outline above will be examined. This could easily be done with large archival datasets and would not require the collection of new data.

Finally, given that learning occurs for students that are nested within classrooms, universities, and countries, future research should examine attempts to examine the size and nature of the classroom, university, or country effect on the relationship between MSLQ constructs and academic performance.

7. Conclusion

The MSLQ is based on important theoretical insights into the nature of learning and the determinants of academic performance. Despite some of the identified psychometric problems, we believe it to be an important tool for educational researchers and practitioners, particularly when paired with measures designed to assess stable student attributes and behaviors.

References


Drasgow, F., Chernyshenko, O. S., & Starke, S. (2010). 75 years after Likert: Thurstone was right. Industrial and Organizational Psychology: Perspectives on Science and Practice (pp. 3).


FURTHER READING


* denotes sources included in the meta-analyses.