The International Journal of Research and Review

Volume 6(2), April 2011

The International Journal of Research and Review is an international interdisciplinary journal that publish empirical reports in the various fields of Social Science.
The International Journal of Research and Review
An interdisciplinary journal on various fields of the Social Sciences

© 2011 Time Taylor International
All rights reserved. No part of this publication may be reproduced or stored in a retrieval system, or transmitted, in any form or by means, electronic, mechanical, photocopying, recording, or otherwise, without permission in writing from the publisher.

The International Journal of Research and Review (TIJRR) is abstracted in the TURR is now abstracted and indexed in the Social Science Research Network, Asian Education Index, Summons Serial Solutions Index by Proquest, Ulrich’s Web, Google Scholar, NewJour, Open J-Gate, Journal Finder, Index Copernicus Journals Master List, SSCI, and EBSCO. TIJRR is also hosted by WT Cox Journal finder, EBSCO, and ERIC on-line.

The TIJRR (ISSN 2094-1420) is published two times a year by Time Taylor International. This journal is part of the Asian EFL Journal on-line services. Access to on-line table of contents and articles are available to all researchers at http://journalofresearchandreview.books.officelive.com/default.aspx for details.

The International Journal of Research and Review (TIJRR) is an international refereed and abstracted journal that publish empirical reports in the various fields of arts, sciences, education, psychology, nursing, social consequences of computer science, and business. Contributors from different institutions are welcome to submit their manuscripts for review and publication on-line that is relevant and significant in the various fields. The International Journal of Research and Review publishes articles and studies on-line twice a year.

Submission Guidelines

The empirical reports featured in the TIJRR are diverse considering the varied fields it can accommodate. The types of empirical reports include:

1. Research Article – ranges from basic and applied research empirical studies employing complex methodologies such as experimentation, survey, evaluation etc. using qualitative or quantitative studies.
2. Literature Review and Metanalysis Studies – synthesis of reviews from journals are viable in this category.
3. Commentary article on theories and models – Issues on previous theories and models are acceptable.

Articles are submitted to the editor at tijrr@yahoo.com. A cover letter indicating contact information of the author(s) is submitted together with the manuscript in word format.

Manuscript Preparation

Submitted manuscript should be typed single spaced. Consult the “Publication Manual of the APA” (latest editions) for detailed guidelines in writing and formatting the manuscript.
The International Journal of Research and Review  
Table of Contents  
Volume 6 Issue 2, April 2011

1   The Computerization of the Self Regulated Learning Assessment System: A Demonstration Program in Developmental Mathematics  
*John Hudesman, Matthew Carson, Bert Flugman, Dorie Clay and Sharlene Isaac*

19  Assessing Cognitive and Metacognitive Learning Strategies in School Children: Construct Validity and Arising Questions  
*Stephan Wernke, Uta Wagener, Andrea Anschuetz, and Barbara Moschner*

39  The Types and Nature of Questions vis-à-vis Students’ Test-Taking Skills as Significant Indicators of Second Language Examinees’ Performance on the TOEFL-ITP Reading Comprehension Sub-Test  
*Analiza Perez-Amurao*

57  The Use of Study Strategies on Mathematical Problem Solving  
*Carlo Magno*

83  Confirmatory Factor Analysis of the Academic Procrastination Scale  
*Romel A. Morales*

Date Published: February 13, 2011
The Computerization of the Self Regulated Learning Assessment System: A Demonstration Program in Developmental Mathematics

John Hudesman  
*SEEK Program, New York City College of Technology and Center for Advanced Study in Education, CUNY Graduate School and University Center*
  
Matthew Carson  
*Dept of Computer Science, CUNY Graduate School and University Center*
  
Bert Flugman  
*Center for Advanced Study in Education, CUNY Graduate School and University Center*
  
Dorie Clay and Sharlene Isaac  
*SEEK Program, New York City College of Technology*

Abstract

The self-regulated learning (SRL) program model is based on a metacognitive (learning how to learn) approach that has been demonstrated to be highly effective in helping students improve their academic performance, with especially impressive results in mathematics. However, SRL program implementation has been hampered by the demands it makes on math instructors, who need to: 1) gather information on students’ quiz scores, 2) calculate the relationship between this information and data on students’ SRL behaviors, and 3) present the results clearly and in ways that will help students strengthen their mathematical understanding and its relationship to SRL behaviors. To make the procedure more efficient, engaging and effective, we have created a first-of-its-kind computerized version of the SRL quiz-taking program using tablet PCs to summarize the data and present a range of math content and SRL information to students. Our results showed that students found the tablet easy to use and described themselves as engaged in the process. This report contains samples of the SRL tablet PC program and student responses.

Keywords: self-regulated learning, metacognition, mathematics achievement

Introduction

It has been repeatedly demonstrated that poor math preparation is a leading cause of academic difficulty for many incoming two-year and technical college students. The Strong American Schools report (2008) found that more than 40% of high school graduates who enter two-year colleges require mathematics remediation at a cost of between 1.85 and 2.35 billion dollars per year. In some community college systems, up to 80% of incoming students require some form of mathematics remediation. Bailey (2009) reports that students can spend up to two years before they have completed the sequence of required developmental mathematics courses, concluding that most of these students are unable to run this gauntlet and fail out of school. Similarly, the Carnegie Foundation (2009) report found that between
60% and 70% percent of these developmental students do not successfully complete the sequence of required courses.

However, passing developmental mathematics is not a reliable indicator of future successful performance. For example, at the technical college of a large university system in the northeast only 55% of students passed a credit-level introductory college-level mathematics course after successfully completing developmental mathematics. As early as the 1990’s Hudesman (1997) studied the relationship between students’ mathematics achievement and academic success by tracking incoming associate degree-students over a six-year period. Prior mathematics achievement, as measured by the student’s score on the mathematics portion of the college’s entrance test, was associated with a variety of both positive and adverse academic outcomes. For example, students who initially passed the mathematics portion of the entrance test had a mean graduation rate of 30% in the School of Technology and Design, whereas students who failed this test had a mean graduation rate of only 13%. Clearly, as expected, proficiency in mathematics represents a major gateway to academic success in two-year and technical college programs.

**Self-Regulated Learning: An Approach that Makes a Difference**

Most interventions designed to address the needs of at-risk students in college developmental courses have focused on teaching academic content, such as mathematics or writing, together with a variety of academic/study skills, such as note-taking and test-taking. However, reviews of intervention studies have revealed that most such programs do not help students to attain their academic goals (Bailey, 2008; Lucas 2007; Simpson, Hyned, Nist, & Burrell, 1997). Often students continue to use maladaptive learning methods because they fail to recognize the constructive potential of the feedback they receive. Incorporating the correct use of feedback into classroom instruction represents a necessary additional component of the learning process. Implementing the Self-regulated learning (SRL) model in the classroom represents such an approach.

Self-Regulated Learning (SRL) involves teaching students and faculty members a new way of understanding the learning process and how to monitor and manage it. Many theories of self-regulation share common elements such as goal setting, the use of strategies, monitoring, and evaluating one’s actions, (Pintrich & Zusho, 2002; Zimmerman, 2002; Butler and Winne, 1995). The SRL theoretical approach described in this article is closely aligned with Zimmerman’s (2002) model which consists of three main sections, forethought, performance, and self-reflection; however, our model is characterized by a feedback cycle consisting of three phases planning, practice, and evaluation (Hudesman, White, & Crosby 2004). The model is represented in more detail in Figure 1.
During the planning phase, students in our SRL programs conduct an academic task analysis, choose those strategies that best address their specific learning challenge, set identifiable goals, and make self-efficacy and self-evaluative judgments. Within the practice phase, students design and carry out their implementation plan. During this phase, they learn to monitor their progress and make appropriate on-the-fly adjustments to their plan. In the evaluation phase, students learn to assess the effectiveness of each intervention. That is, they use self- and instruction-generated feedback to build on successful strategies and modify or replace less effective strategies. The students’ responses from the evaluation phase then become the basis for the planning phase in the next SRL cycle. Implicit in the SRL model is that learning involves a series of these learning cycles each of which brings the student closer to his/her academic goal. This cyclical model provides students.
with multiple opportunities to constructively make use of the instructor's feedback. A number of investigators have demonstrated that students trained in self-regulation can demonstrate improvement in their academic achievement (Schunk 1996; Paris & Paris 2001; Cleary & Zimmerman 2004). Additionally, a meta-analysis of studies of elementary and high school students found that self-regulation produced statistically large achievement effects, (Dignath & Buettner, 2008). These gains were especially noteworthy in the area of mathematics.

In working with instructors and counselors to implement SRL based programs, we have been guided by two general principles: 1. Instructors and counselors must give students fast, clear, and accurate feedback about their performance, and 2. Instructors must insist that students demonstrate that they understand this feedback and can use it to improve their performance. These implementation principles are very similar to those emphasized in formative assessment programs that have become a priority for college administrators, accrediting agencies and educational investigators.

Self-regulation increases the likelihood that students are more likely to take responsibility for their learning. Self-regulated students understand that academic success is a function of experimenting with different strategies and not a function of ‘natural intelligence’ or some other external force such as whether the instructor likes them or not (Zimmerman, 2002). The power of the SRL model is highlighted in a classic study by Zimmerman and Bandura (1994) who demonstrated that SRL skills are more highly correlated with college grade point average than are scores on the Scholastic Aptitude test (SAT).

Self-Regulation: What We Know about the Application of an SRL Program Model.

Over the last 10 years, our SRL program group has iteratively developed and researched various components of an SRL Program model. A variety of SRL based programs have been designed and implemented in high schools and colleges in New York, New Jersey, and Ohio (Blank, Hudesman, and Zimmerman, 2007; Hudesman, 2005, 2010; Zimmerman, Moylan, Hudesman, White, and Flugman in press). Many of these initiatives have been funded by major federal agencies including the Institute for Education Sciences (IES), the Fund for the Improvement of Post Secondary Education (FIPSE), and the National Science Foundation (NSF). The IES study described below has particular relevance to our present demonstration program.
The Application of the Model in Developmental Mathematics: An IES Development Study Program.

Zimmerman, Moylan, Hudesman, Flugman, & White (in press) developed a paper-and-pencil SRL semester-long classroom intervention and then pilot tested its effectiveness with incoming associate degree students enrolled in developmental mathematics courses. All of the mathematics sections were taught by experienced instructors. Two hundred and eight students were randomly assigned to either an experimental classroom (receiving the self-regulatory intervention) or a control classroom (receiving conventional instruction) for a 15-week semester. The program procedure, which is also used in the present demonstration project, consisted of:

1. A series of math quizzes. Every two to three class sessions, students in the SRL sections were administered a 15-20 minute quiz involving four mathematics problems as a vehicle for frequent feedback to students and teachers (Appendix A). These quizzes were formatted so that both before and after attempting to solve each problem students were required to make confidence judgments indicating how sure they were that they could correctly solve the math question, i.e., they were asked to make task-specific self-efficacy judgments before solving individual problems and self-evaluative judgments after attempting to solve each math problem. The rational for this process is that many students ‘don’t know what they don’t know’, and they consistently overestimate their self-efficacy and self-evaluative judgments. By making these judgments, and then receiving feedback about their accuracy, students become more accurate in calibrating these critical self-regulation processes and apply this knowledge as part of the process of selecting appropriate strategies. For example, students who think they understand the material, when in fact they do not, are less likely to prepare for their next quiz.

2. The self-reflection process. After receiving graded quizzes from the instructor, SRL students had the opportunity to earn additional quiz credit by completing self-reflection and mastery-learning form designed to guide student’ thinking about their erroneous answers to items on the quiz. The self-reflection form (Appendix B) required students to compare their self-efficacy and self-evaluative judgments with their actual performance on the quiz item, explain their ineffectual strategies with regard to solving the mathematics quiz item and establish and try out new, more effective mathematics and SRL strategies. A main portion of the reflection form requires that students redo the incorrect quiz question and include a description (in words) of the step-by-step strategies that were used to solve the problem. Students also had to solve a new problem that required them to use similar strategies.

3. Program results. Although the students in both the SRL and control group developmental math students had similar pre scores on the
mathematics portion of the American College Testing program ACT, i.e., the COMPASS (2006), SRL students demonstrated greater academic progress than did students enrolled in the control group sections. Indications of progress include: 1. higher mean scores on major examinations given periodically during the semester: 69.98 compared with 63.12 (ANOVA=4.6, p<.05, effect size =.37) for the students who completed the semester in the SRL and control group sections respectively; 2. Higher pass rates on a departmental final examination, 52% compared with 31% (Chi square = 9.42, p<.01, effect size =.44) for SRL and control group students respectively. Finally, 46% of the students initially enrolled in the SRL sections of developmental mathematics passed the COMPASS compared with only 25% of the students enrolled in the control group sections (Chi square= 9.94, p<.01, effect size=.45, cohort analysis). The COMPASS outcomes become more striking if we compare only those students who completed the course, i.e., did not withdraw during the semester (N=140). Sixty-four percent of the students enrolled in the SRL sections who completed the course passed the COMPASS vs. only 39% of the students enrolled in the control group sections (Chi square = 8.13, p<.01, effect size =.50).

**Computerizing the SRL Assessment System**

This study, together with a number of other research and development SRL programs (Blank, Hudesman & Zimmerman, 2007; Hudesman, Zimmerman, & Flugman 2010), demonstrates that while using this program model can significantly improve student performance. However, there were a number of issues that emerged during the program implementation. These issues, and how they can be effectively addressed by the computerized version of the SRL Assessment System, include the following:

1. When using the paper and pencil version of the SRL Assessment Program students only completed only about two-thirds of the self-efficacy and self-evaluative judgments on the quizzes. We know that completing these judgments is associated with an improved understanding of their self regulation processes (Zimmerman, Moylan, Hudesman, White, & Flugman, in press). The tablet PC version of the program requires students to take their quiz and show their work on the tablet. As part of this process students must make 100% of their SRL judgments because they are not able to continue on to the next quiz question until all the SRL judgments have been made.

2. Instructors often expressed concern about having to discuss the relationship between the self-efficacy and self-evaluative judgments and students’ quiz performance indicating that they are math teachers and not educational psychologists. The tablet PC model automatically addresses this problem by creating summary graphics of the self-efficacy and self-evaluative judgments and quiz scores for each student.
and for the entire class, thus making it much easier for instructors and students to understand the relationship between SRL processes and mathematics performance.

3. Instructors indicated that implementing the entire SRL program takes too much time. The tablet PC version of the program automatically provides each student’s quiz scores and allows the instructors to make constructive comments on the electronic copies of the students’ quizzes. It also enables the instructor to store assessment material more efficiently.

The computerization of the SRL assessment process that is demonstrated in this program is the first-of-its-kind. It is designed to enhance the effectiveness of the SRL assessment and instruction intervention. Using the tablet PC is particularly well suited to Science, Technology, Engineering, and Mathematics (STEM) students in general and mathematics students in particular because it allows them to show all of their work, including formulas, symbols, and the like. It would be very cumbersome for students to show their work on a regular computer as they would have to hunt and peck for each symbol, a process that would soon become impractical. Within this context, this demonstration program was specifically designed to address the following questions:

1. Can we create a computerized version of the successful paper and pencil SRL Assessment System that used in the IES study described above?
2. Will students find this system acceptable and be willing to use it on an ongoing basis in the classroom?
3. Will the graphical representations depicting the relationship between their mathematical and metacognitive processes be helpful to students?
4. Does the tablet PC version of the SRL Assessment System, in comparison to the paper and pencil version, facilitate the academic progress of students in a developmental mathematics course?

Method

Participants

Students were registered in a six week summer session developmental mathematics course at the technical college of a large northeastern university. The course was delivered to all students using the SRL method previously described. The course had an enrollment of 18 entering freshmen all of whom had failed the mathematics portion of the American College Test, i.e., the COMPASS (2006). Nine students volunteered to use the tablet PC version of the SRL Assessment System and nine decided to use the
traditional paper and pencil version of the SRL program. The course was taught by an experienced developmental mathematics instructor. As is the case for all of the college’s summer developmental mathematics courses, there was an in class tutor.

**Materials and Procedure**

As part of the course, the instructor administered a series of seven short quizzes, none of which exceeded 20 minutes. Each quiz contained four questions, each with a value of 25 points. Starting with the fourth quiz, the tablet PC program students took their quizzes on an HP TouchSmart tm2 tablet PC. The format for using the tablet PC was for students to log in and then take the quiz, one question per screen. In keeping with the SRL program, described in an earlier section of this report, the students were instructed to read each question, make a self-efficacy judgment before answering the question, answer the question showing all their work, and then make a choice of the correct answer from among five choices. Students then made a self-evaluative judgment, after which the next question would appear. Students were unable to proceed to the next question without making the self-evaluative judgment for the previous question. At the end of the quiz students were asked to answer two questions about their ‘time on task,’ i.e. how much time they spent on homework, and how much time they spent preparing for the quiz.’ The student’s work was then uploaded to the instructor’s tablet PC as well as to the college’s server.

Uploaded work was automatically scored by the computer which generated the following information that was immediately available to the instructor, tutor, and the student: 1. Score on the quiz; 2. an indication for each question as to whether it was correct or incorrect; 3. A series of graphs that related the students’ quiz score to their SRL behaviors, i.e., their self-efficacy and self-evaluative scores, as well as to their time on task. These graphs were cumulative in nature thus allowing students and instructors to graphically track the relationship between the math and SRL behaviors over time.

The instructors was also able to view each student’s uploaded quiz on her tablet PC, make written comments directly on her tablet PC, award partial credit, and send a copy of the entire package to the student’s home computer as a PDF. Students were then able to use this information to make appropriate revisions on incorrectly answered questions by using the self reflection (mastery learning) form (Appendix B). Based on the quality of the students’ work it was possible for them to earn up to 100% of the credit value of the original incorrectly answered question.

As part of this demonstration program, those students using the tablet PCs were asked to complete a total of three, three-question surveys on how using the tablet PC compared with taking quizzes using a paper and pencil.
format. An additional survey, administered to all students during the last week of the course, asked for their reactions to various SRL interventions, including the use of tablet PCs.

Results

Individual Student Results

As previously mentioned, tablet PC users received an email packet after each quiz. The packet consisted of the corrected quiz and two graphs. Students used the tablet PCs for quizzes 4 – 7. Based on the information supplied by the student, the computer automatically created graphs that illustrate the relationship between the student’s math and SRL skill levels. Sample graphs for two program students are presented. Figure 2 shows the relationship between the student’s actual quiz score and his SRL judgments, i.e., his self-efficacy and self-evaluation judgments.

![Math Quiz Scores and SRL Judgments](image)

**Figure 2.** Results of the Relationship between One Student’s Math Quiz Scores and SRL Judgments. Notes: 1 Mean pre self-efficacy and post self-evaluative judgments were transformed from a scale of 1-5 to a scale of 20 – 100; 2. This student was absent for quiz 6.

The score is the student’s actual score for each of the quizzes. The self-efficacy judgment (AvgPre) represents the mean of the student’s self efficacy judgments for each quiz. Similarly, the self-evaluation judgment (AvgPost) is the mean of the student’s self-evaluation judgments for each quiz. What is noteworthy in this graph is that at the start of the process (quiz 4) the student’s SRL judgments are noticeably higher than his actual score. This discrepancy illustrates the difference between what the student thinks he knows in terms of his judgments (a lot) and what he actually knows in terms
of his quiz score (not so much). However, by the end of the seventh quiz the gap has narrowed considerably indicating that what the student thinks he knows is much more closely aligned with what he actually knows. This self knowledge is critical for the student if he is to develop more adaptive academic behaviors.

Figure 3 illustrates the relationship between another student’s time-on-task on each quiz, as measured by time spent on homework as well as time spent in quiz preparation, and his quiz score.

The two time-on-task measures were calculated on the basis of how many minutes students spent studying by: 1. doing homework problems (Time HW), and 2. the number of minutes that they spent studying for the quiz in addition to their homework. Time on task was calculated by categorizing the number of minutes as follows: 0 minutes = 0, 1 – 29 minutes = 25, 30 – 59 minutes = 50, 60 – 119 minutes = 75, 120 minutes (+) = 100. What is noticeable is that when the student’s preparation time decreases (see quiz 6) so does his quiz score. Conversely, when the student’s preparation time increases so does his quiz score (see quiz 7). Providing this type of feedback to students on an ongoing basis can be an important feedback aid in assisting students to see the relationship between effort and results.

Class-wide Responses to the Project

On three occasions students using the tablet PCs were given a brief survey. Seven to eight students out of the nine participants responded to each
question. Students were asked: 1. to compare taking the class quizzes using
the tablet PC with taking the quizzes using a paper-and-pencil format, 2.
whether it was helpful to receive computer generated graphs that related
their math outcomes to SRL behaviors, and 3. whether there was a certain
‘cool factor’ to using the tablet PC. Table 1 summarizes the cumulative
responses for the three surveys.

Table 1
Survey Responses for Students Who Used the Tablet PC

<table>
<thead>
<tr>
<th>Measures</th>
<th>Cumulative Survey Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Is the tablet pc-paper easier than pencil paper?</td>
<td>9</td>
</tr>
<tr>
<td>Are Graphs helpful?</td>
<td>20</td>
</tr>
<tr>
<td>Is there a ‘cool factor’ using the tablet pc?</td>
<td>20</td>
</tr>
</tbody>
</table>

Notes. (1) There were seven respondents in the first and second surveys, and eight respondents in the third survey. (2) NA indicates that there was no ‘same answer’ option applicable for this question.

As can be seen by the response pattern in Table 1, all of the students
who used the tablet PCs found it at least as easy to use as taking a paper-
and-pencil quiz. Almost all of the survey respondents found it helpful to
receive the tablet PC graphs that illustrate the relationship between their
mathematics performance and their SRL behaviors. This type of graphical
feedback is not available to students (or instructors) using the paper and
pencil version of the SRL Assessment Program. And almost all of the
students who used a tablet PC thought there was a ‘cool factor’ in using this
technology.

In addition to the three brief surveys completed by the tablet PC users,
all the students (tablet PC users and non-users) were surveyed at the end of
the course about their reactions to a variety of SRL course interventions,
including the use of tablet PCs. The results of this survey are summarized in
Table 2.

As seen in Table 2 almost all students in both groups responded
positively to a wide variety of SRL course features including the use of tablet
PCs. All of the students in the class, except for one student who did not use
the tablet PC, recommended that tablet PCs should be used in future
mathematics classes.

These survey results reflect the reports of informal discussions that
took place between students, the instructor, the tutor and which indicated
that the tablet PC information was quite helpful.

To determine the extent to which using the tablet PC would in any way
show promise in facilitating students’ academic progress in the course, tablet
PC students and paper-and-pencil students were compared on a number of
measures.
Table 2
**Survey Responses for All Students Taking the SRL Developmental Mathematics Class**

<table>
<thead>
<tr>
<th>Questions</th>
<th>Tablet Users (N=9)</th>
<th>Non Tablet Users (N=9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was it helpful to have frequent quizzes?</td>
<td>Yes: 9</td>
<td>No: -</td>
</tr>
<tr>
<td>Was it helpful to make confidence judgments before and after each quiz?</td>
<td>Yes: 7</td>
<td>No: 2</td>
</tr>
<tr>
<td>Was it helpful to make grade predictions?</td>
<td>Yes: 7</td>
<td>No: 2</td>
</tr>
<tr>
<td>Was it helpful to enter your study/hw prep time?</td>
<td>Yes: 8</td>
<td>No: 1</td>
</tr>
<tr>
<td>Was it helpful to complete self–reflection forms for incorrectly answered questions?</td>
<td>Yes: 9</td>
<td>No: -</td>
</tr>
<tr>
<td>Is there a connection between math and non math (SRL) activities in this class?</td>
<td>Yes: 9</td>
<td>No: -</td>
</tr>
<tr>
<td>Would you recommend the tablet pc be used in future classes?</td>
<td>Yes: 9</td>
<td>No: 1</td>
</tr>
</tbody>
</table>

Table 3
**A Comparison of Mean Course Measures and COMPASS Scores for Tablet PC Users and Non Users**

<table>
<thead>
<tr>
<th>Course Measures</th>
<th>Tablet PC Users (N=9)</th>
<th>Non Tablet PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre Course COMPASS: Algebra Score</td>
<td>24.00</td>
<td>23.89</td>
</tr>
<tr>
<td>Mean Number of Absences</td>
<td>1.44</td>
<td>2.11</td>
</tr>
<tr>
<td>Mean Homework Score</td>
<td>62.22</td>
<td>46.78</td>
</tr>
<tr>
<td>Mean Quiz Score</td>
<td>87.82</td>
<td>74.85</td>
</tr>
<tr>
<td>Mean Test Score</td>
<td>79.33</td>
<td>70.78</td>
</tr>
<tr>
<td>Final Exam</td>
<td>82.44</td>
<td>76.67</td>
</tr>
<tr>
<td>Course Average</td>
<td>81.56</td>
<td>74.56</td>
</tr>
<tr>
<td>Post Compass: Algebra Score</td>
<td>35.38</td>
<td>36.63</td>
</tr>
</tbody>
</table>

While it was not the intent of this demonstration to determine the relative effects of tablet PC vs. paper and pencil versions of the program on academic performance, we were interested in whether there would be any consequences of using technology in what was already a successful program. The data in Table 3 indicate that while the tablet PCs were only used by students for approximately half the course, the tablet PC group performed slightly better than students in the paper and pencil group on six out of seven academic measures (absences, homework scores, quiz scores, periodic exam scores, final examination score, and course average). Obviously, a larger more rigorous efficacy study needs to be done in order to investigate the potential
of the SRL tablet PC system to assist students to improve their academic performance over and above the pencil-and-paper version of the program.

**Discussion**

Multi-component instructional systems that emphasize ‘learning how to learn’ concepts and skills are often difficult for instructors to deliver in content area classes resulting in their non- or underutilization. In the latter case, students and instructors not only have difficulty in managing all of the parts of the program but also struggle to fully understand metacognitive concepts and how they can be used and learned.

The purpose of this demonstration program was to address this overall challenge by designing and implementing a computerized version of the SRL Assessment System that built on the earlier successes of the paper-and-pencil version. Our goal was to optimize the delivery of components, make it more efficient, and determine if these changes showed promise in effecting student outcomes. The computerization was made all the more difficult by the need to capture student work in mathematics, including symbols and equations, necessitating the use of tablet PCs.

While only one component of the SRL system was computerized and used for only approximately half the course, the results indicated that computerization proved feasible for classroom use, was perceived as useful and interesting by the students who used it as well as those students who observed its use, and showed preliminary promise of improving student outcomes. Specifically, our results showed that computerization was able to address problems that emerged during the implementation of the paper-and-pencil version of the program in that: 1. Students generated all the self-efficacy and self-evaluation data, making it possible for the instructor to do a more complete analysis of their quiz scores. 2. Students in the tablet PC group (and the instructor) received immediate feedback regarding quiz scores which could be acted on while the pencil and paper students needed to wait for their quizzes to be hand scored. 3. Students in the tablet PC group were able to receive an email PDF package that included instructor comments and suggestions regarding their work as well as and graphs illustrating the relationships between their mathematics quiz scores, self-efficacy and self-evaluation judgments, and time-on-tasks; 4. The tablet PC automatically saved all of the students work. By giving instructors access to this type of stored data they have the ability to analyze student performance on an individual or group level for formative assessment purposes. Although a potentially powerful process, formative assessment is frequently limited by the instructor’s inability to store and manage large amounts of student data (Heritage, 2010). In terms of outcomes, there were some indications that students in the tablet PC group demonstrated better performance on a number of course related academic measures, e.g., the final examination.
These conclusions are limited by the small number of students enrolled in the demonstration program and by its implementation in one class for approximately half of the course. Despite these limitations, we were encouraged by the observation that our results consistently favored the tablet PC group. Future research needs to increase the number of students and instructors who implement the program and evaluate it using an experimental design. To date, however, we have begun the computerization of a complex intervention that has the potential to optimize its implementation in a wide range of settings.

References


**Author Notes**

The authors would like to express their gratitude to Grazyna Niezgoda who was the course instructor.

This program was funded by the CUNY Graduate School and University Center and the SEEK Program, New York City College of Technology.
Appendix A

Quiz #

Professor

Name: ___________________________  Date: __________

**Before solving each problem, circle the number that represents how confident you are that you can solve it correctly.**

**REMEMBER!**

Show all your work.
Simplify all your answers.

**After you have solved each problem, circle the number that represents how confident you are that you solved it correctly.**

<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4.5</td>
<td>1.2.3.4.5</td>
</tr>
<tr>
<td>1.2.3.4.5</td>
<td>1.2.3.4.5</td>
</tr>
<tr>
<td>1.2.3.4.5</td>
<td>1.2.3.4.5</td>
</tr>
<tr>
<td>1.2.3.4.5</td>
<td>1.2.3.4.5</td>
</tr>
<tr>
<td>1.2.3.4.5</td>
<td>1.2.3.4.5</td>
</tr>
</tbody>
</table>
Appendix B

SRL Math Revision Sheet, Quiz #____ Item # ____
Student: ________________________ Date: ________

Instructor: ________________________ 8pts

Now that you have received your corrected quiz, you have the opportunity to improve your score. Complete all sections thoroughly and thoughtfully. Use a separate revision sheet for each new problem.

PLAN IT

1. a. How much time did you spend studying for this topic area? _______

   b. How many practice problems did you do in this topic area in preparation for this quiz?

      (circle one)  0 – 5 / 5 – 10 / 10+

   c. What did you do to prepare for this quiz? (use study strategy list to answer this question)

2. After you solved this problem, was your confidence rating too high (i.e. 4 or 5)? yes no

3. Explain what strategies or processes went wrong on the quiz problem.

   8pts
PRACTICE IT

4. Now re-do the original quiz problem and write the strategy you are using on the right.

5. How confident are you now that you can correctly solve this similar item?

   Definitely not  Not confident  Undecided  Confident  Very confident

   confident

   1  2  3  4  5

6. Now use the strategy to solve the alternative problem.

   4pts

7. How confident are you now that you can correctly solve a similar problem on a quiz or test in the future?

   1  2  3  4  5
Assessing Cognitive and Metacognitive Learning Strategies in School Children: Construct Validity and Arising Questions

Stephan Wernke  
*University of Oldenburg, Germany*

Uta Wagener  
*University of Osnabrueck, Germany*

Andrea Anschuetz  
*University of Oldenburg, Germany*

Barbara Moschner  
*University of Oldenburg, Germany*

**Abstract**

Self-regulated learning is a complex construct that has inspired innumerable research in recent years. The self-regulated activity of a learner and the awareness in a learning process are shaped by cognitive and metacognitive strategies students use. This study addresses the questions how metacognitive and cognitive learning strategies can be assessed and which different groups of strategies can be identified empirically. For this reason two studies were conducted with a self-report questionnaire. In study one learning strategies in elementary school children (N= 1083) were assessed; in the second study students of the fifth and sixth grade (N= 1067) of a secondary school were tested with the same instrument. The results of both studies show whereas for the cognitive subscales three different dimensions are clearly empirically separable, metacognitive processes in self-regulated learning are more interwoven. Our data reveals that only Planning as a metacognitive strategy can be separated from other metacognitive processes.

**Keywords:** Self-regulated learning, self-report measurements, learning strategies

**Introduction**

Self-regulated learning is a complex construct that has inspired innumerable research projects in recent years. It has been used in different theoretical traditions and examined with different methods. Even if there is no simple, consensual definition of self-regulated learning (cf. Boekaerts & Corno, 2005), there are some common assumptions: Strategic action, metacognition, and (intrinsic) motivation are considered to play a part in a learning process that can be labeled self-regulated learning (e.g. Boekaerts, 1999; Moschner, 2007; Winne & Perry, 2000). The activity and reflexivity of a learner are important aspects: “students are self-regulated to the degree that they are metacognitively, motivationally, and behaviorally active participants in their own learning process” (Zimmerman, 2001, p. 5). This activity of a learner and the awareness in a learning process are shaped by cognitive and metacognitive strategies students use. “In the last decade it has become clear that one of the key issues in self-regulated learning is the students’
ability to select, combine, and coordinate cognitive strategies in an effective way” (Boekaerts, 1999, p. 447).

Cognitive and metacognitive strategies are therefore seen as important aspect of self-regulated learning. How can strategies be assessed and which different groups of strategies can be identified empirically? Does this differ in different age groups? These questions will be addressed in the following article at first theoretically and then empirically.

Cognitive and Metacognitive Strategies in Different Models of Self-regulated Learning

Cognitive and metacognitive strategies are mentioned in every model of self-regulated learning but they are given varying importance. While some approaches distinguish different strategy groups they are more fused in other models. In the following, five different approaches to strategic action and self-regulated learning will be shortly presented to identify the role cognitive and metacognitive strategies play in each of these models.

Research on strategic action has a long tradition in educational psychology. Weinstein and Mayer (1986) differentiate between cognitive, metacognitive, and as a third group motivational and affective strategies. Cognitive strategies include rehearsal strategies, elaboration strategies, and organization strategies. Metacognitive strategies are characterized as comprehension monitoring strategies but are not divided into further subgroups (Weinstein & Mayer, 1986, p. 316).

In contrast to this, Boekaerts (1999) developed a three-layered model of self-regulated learning to capture different areas of regulation. The first layer contains the regulation of processing modes, which includes the choice and application of cognitive strategies. The second layer is the regulation of learning processes and implies the use of metacognitive knowledge and metacognitive skills to direct one’s own learning. The third layer consists of the regulation of self and concerns more general aspects like choice of goals and resources.

Some other models have developed a more detailed description of cognitive and metacognitive processes that are involved. Models of information processing regard cognitive processes with complex feedback loops as the basis of self-regulated learning (Winne & Hadwin, 1998; Winne & Perry, 2000; Zimmerman, 2001). Different processes are distinguished according to their chronology in the learning episode which is conceptualized as information processing. Defining the task (1), setting goals and planning how to reach them (2), enacting tactics (3), and adapting metacognition (4) are the four phases that are separated by Winne and Hadwin (1998). Metacognitive monitoring and metacognitive control are distinguished as two events that are relevant in each of these phases.

Zimmerman (2000, 2001) postulates three phases, the forethought phase, the performance or volitional control phase, and the self-reflection phase. He distinguishes task analysis including goal setting and strategic
planning in phase one, self-control (volitional control) and task-related strategies in phase two, and self-reflection and self-evaluation in phase three. In each of these phases, different metacognitive processes are relevant and different strategies can be applied for planning, controlling, and evaluating the learning process.

Pintrich (2000) has also developed a temporal model of the process of self-regulated learning in which four phases are distinguished. In his conceptualization the first phase is called forethought, planning and activation including goal setting. The second phase comprises the monitoring of the learning process. The third phase includes regulation and control, thus the use of control strategies is part of this phase. The fourth phase is called reaction and reflection and consists of all evaluations, judgments, and attributions that are made subsequently to a learning episode. Pintrich (2000) points out that the described phases represent a time-ordered sequence. However, all phases do not take place in every learning process and they do not always happen consecutively. According to Pintrich (2000), the four phases of self-regulated learning can occur in four different areas: cognition, motivation, behavior, and context. It is important to note that phases and areas of regulation are not necessarily independent and distinct. “The phases may overlap, occur simultaneously with multiple interactions among the different processes and components” (Pintrich, 2000, p. 456). Again, different strategies are to be applied in different chronological phases of the learning process.

Metacognitive strategies of planning, monitoring, and evaluating are relevant as well as different cognitive strategies for dealing with a complex learning content.

Cognitive and metacognitive strategies are distinguished in these models but according to Veenman, Van Hout-Wouters, and Afflerbach (2006) the relationship between cognition and metacognition and between cognitive and metacognitive strategies is complex and not easy to disentangle. Metacognition is regarded as higher-order cognition about cognition. But at the same time metacognition is also cognition. Veenman et al. (2006) point out that metacognition is contingent on cognition. Domain-specific knowledge is necessary to apply metacognitive functions adequately. Metacognitive planning cannot be used without cognitive activities referring to the task at hand.

Developmental Aspects

Developmental aspects are crucial if we focus on self-regulated learning in school children. Using different research methods and different theoretical backgrounds the beginning of the use of cognitive and metacognitive strategies has been located in very different age groups from pre-school children to adolescents. However, it is beyond doubt that there is a considerable development from primary school children to adolescents. For cognitive as well as metacognitive strategies it can be summarized that their use advances with age which does not only concern
the quantity of strategy use but also the quality (e.g. Artelt, 2006, Hasselhorn, 2004).

Around the age of ten a quantitative increase in the use of elaboration strategies as well as an improvement in the application of strategies has been reported (Artelt, 2006). Hasselhorn (2004) argues that children around the age of ten are beginning to be able to reflect about their own abilities, their own learning, and their knowledge in a more abstract manner and that this is the basis of metacognitive processing. According to Hasselhorn (2004), at this age cognitive learning strategies are also used spontaneously for the first time (cf. Hasselhorn, 1992). Between the age of eleven to twelve a considerable increase in cognitive and metacognitive processes is described, which is becoming more differentiated and more effective in twelve to sixteen year old adolescents (Hasselhorn, 2004).

**Self-report Measures of Self-regulated Learning**

While there are several possibilities to measure self-regulated learning, questionnaires using self-report items are the most frequently used instruments in large samples. The two predominating measures in Anglo-American studies which are based on cognitive theories are the “Learning and Study Strategies Inventory” (LASSI; Weinstein, Palmer, & Schulte, 1987) and the “Motivated Strategies for Learning Questionnaire” (MSLQ; Pintrich, Smith, Garcia & McKeachie, 1993). In Germany the most frequently used questionnaires are two instruments: the “LIST” (Lernstrategien im Studium; Wild & Schiefele, 1994) is constructed for the assessment of learning strategies in higher education, the “KSI” (Kieler Lernstrategieinventar; Baumert, Heyn & Koeller, 1992) is appropriate for the assessment of learning strategies of school children. In the following these instruments and results of the dimensional structure of cognitive and metacognitive strategies are described briefly to illustrate theoretical assumptions in questionnaire construction and empirical findings in research.

**LASSI - Learning and Study Strategies Inventory**

Weinstein und Mayer (1986) differentiate three dimensions of learning strategies: cognitive, metacognitive, and affective strategies. The LASSI divides affective strategies into the six subscales Concentrating, Anxiety, Motivation, Attitude, Time-Management, and Test-Strategies. Cognitive strategies are operationalized by the three subscales Selecting the main ideas, Information Processing, and Study Aids. The scale Self-Testing includes metacognitive strategies. Reported reliabilities in the User’s Manual (Weinstein et al., 1987) are between α=.68 and α=.86. The authors themselves did not check the construct validity of the LASSI. Factor analyses of subsequent studies with the LASSI (e.g. Olaussen & Braten, 1998) did not prove the postulated structure; only three factors were distinguished. Factor 1 includes “Effort-related Activities”
(Motivation, Time Management, Concentration), factor 2 „Goal Orientation“ (Information Processing, Study Aids, Self-Testing), and factor 3 „Cognitive Activities“ (Selecting Main Ideas, Test Strategies, Anxiety).

**MSLQ - Motivated Strategies for Learning Questionnaire**

Pintrich and his colleagues (1993) developed another questionnaire based on the conception of Weinstein et al. The MSLQ (Pintrich et al., 1993) includes two main sections: Motivation on the one hand and learning strategies on the other. The learning strategies scales are divided into three categories: The use of metacognitive and cognitive strategies and the management of different learning resources. Cognitive strategies are separated into Rehearsal, Elaboration, Critical Thinking, and Organization. Subscales of the metacognitive strategies are Planning, Monitoring, and Regulation. The subscales measuring the Resource Management are Time Management, Study Environment, Effort Management, Peer Learning, and Help Seeking. The reported reliabilities of the scales are between $\alpha=.52$ and $\alpha=.80$. The factorial structure of the MSLQ was proved in several studies. Only the structure of the metacognitive strategies could not be differentiated (Garcia & Pintrich, 1996).

**LIST – Lernstrategien im Studium**

The LIST questionnaire (Wild & Schiefele, 1994) is a German adaptation of the MSLQ. It includes the same scales and subscales as the original instrument. The authors report reliabilities between $\alpha=.64$ and $\alpha=.90$. Factor analyses show comparable results as were shown in studies with the MSLQ. Most subscales could be replicated, again only the metacognitive strategies could not be empirically differentiated.

**GSSS - Goals and Strategies for Studying Science**

Nolen and Haladyna (1990) developed an instrument for a younger target group (class 9 to 12). The GSSS consists of four scales: Three scales in order to assess the cognitive strategies Memorization, Elaboration, and Organization, and the scale Monitoring to measure metacognitive strategies. The assumed structure was proved in confirmatory and exploratory factor analyses. Reliabilities of the scales are between $\alpha=.64$ and $\alpha=.81$.

**KSI - Kieler Lernstrategie-Inventar**

Based on the instruments MSLQ, LASSI, and GSSS Baumert, Heyn, and Koeller (1992) developed six scales for their KSI. They constructed three subscales for measuring cognitive strategies (Memorization, Elaboration, Transformation) and three metacognitive
subscales (Planning, Monitoring, Regulation). The reliabilities of the scales are between $\alpha=.77$ and $\alpha=.92$ (Baumert, 1993). The structure of the subscales was supported by a confirmatory factor analysis; however, subsequent studies (e.g., Spoerer, 2004) could not differentiate the subscales of the metacognitive strategies.

To sum up, it is obvious that there are several reliable questionnaires for measuring learning strategies. But regarding the construct validity there are some open questions. On the one hand, the structure of cognitive strategies is empirically proved and replicated several times. On the other hand, the structure of the metacognitive strategies is still unclear. In most studies, the subscales could not be differentiated.

**Method Study 1**

**Participants**

Participants of this study were $N=1083$ German fourth grade elementary school children. There were 540 girls (49.9%) and 542 boys (50.0%) children (1 missing). The age group ranged from eight to twelve years, and the average age was 9.7 years ($SD=.63$).

**Instrument**

The questionnaire used in this study was developed to assess the use of learning strategies of elementary school children. Its structure is based on existing instruments and consists of seven learning strategy scales and 46 items. The items have to be answered on a four-point likert-scale ranging from “not true” to “totally true” which is illustrated for the children with smilies. None of the participants expressed difficulties to understand the answering format. All categories were chosen.

In contrast to common self-report questionnaires this instrument aims at measuring the use of learning strategies as close to the learning process as possible. Therefore, the respondents are confronted with a typical school situation at first. They have to read a non-fictional text and after that they have to answer some questions about it. Directly after this, when the learning process is still tangible, the respondents have to report their use of learning strategies. All items refer to this learning process explicitly, e.g. “While reading I asked myself if I had understood everything”. The text (“Nights of the Pufflings”) and the questions about the text were taken from the Progress in International Reading Literacy Study (PIRLS) (Mullis et al., 2003). They were designed according to the different competence levels of the PIRLS study. The curricular validity of this text was examined by international reading and curriculum experts. Some of the questions have to be answered in multiple choice format and others in an open answering format.

The seven scales of the instrument refer to cognitive and metacognitive learning strategies. Cognitive strategies include students’
use of basic and complex strategies for the processing of information from texts and lectures (cf. Weinstein & Mayer, 1986). For assessing the cognitive learning strategies three subscales were designed.

Elaboration refers to the strategies students employ to understand information by activating their prior knowledge and creating analogies (8 items).

Rehearsal refers to strategies of memorization and involving rehearsing of items in order to remember them (7 items).

Organization refers to the strategies students employ to select information and construct connections (6 items).

For assessing metacognitive learning strategies four subscales were constructed (cf. Pintrich, 2000).

Planning involves setting educational as well as task analysis (7 items).

Monitoring strategies are used in order to observe the effectiveness of the learning performance and behavior. The scale includes strategies like focusing the attention, comprehension monitoring, and time management (6 items).

Regulation strategies are important if problems appear during the learning process. If learners have comprehension problems or use inadequate learning strategies, they should regulate their learning activities (6 items).

Evaluation strategies are used to check if planned educational goals are reached or not. Students control their own effort and draw conclusions for the next learning episode (6 items).

**Procedures**

The questionnaire was administered in school in regular classrooms by trained staff members of the University of Oldenburg (Germany). The children were informed orally that they were participating in a survey about the way children learn. Additionally, it was pointed out that it was not an achievement test that they would not get any marks, that their answers were handled anonymously, and that their teachers would not get an insight into their answers. Finally, the researcher stressed the importance of being honest.

After handing out the questionnaires the children had the opportunity to ask comprehension questions. Answering the likert scale was illustrated with an item example. There was a time limit of one school lesson (45 minutes) for the completion of the whole survey which includes reading the non-fictional text, answering the questions concerning the text, and answering the questionnaire items. In most cases less time was required, generally between 30 and 35 minutes.

**Research Question**

Referring to the results of research in self-regulated learning, especially with self-report instruments, we ask if the theoretically
assumed and model-based multidimensionality in cognitive and metacognitive strategies can be found in empirical data. Combining this methodological research question with results of developmental aspects we want to investigate if the different dimensions of cognitive and metacognitive learning strategies can be empirically and statistically distinguished already in elementary school children.

Results Study 1

Construct Validation: Factor Analysis

According to the theoretical assumptions and the multidimensionality of self-regulated learning an exploratory factor analysis using principal component factoring with varimax rotation was conducted among the items assessing cognitive strategies and among the items measuring metacognitive strategies. In the factor extraction three procedures were used to identify the underlying factor structure: the Kaiser-Meyer-Olkin measure of sampling adequacy, the Kaiser-Guttman criteria (eigenvalues greater than one), and the scree plot by Catell (Field, 2009). Using all these methods and criteria for extracting the factor structure should reduce the risk of over or under extraction. The varimax rotation method was applied because it accounted for larger factor loadings under each of the factors that will be extracted. Analyzing the factor structure, there were different criteria defined in advance: items should be assigned to factors based on their factor loadings, items with factor loadings below .45 should be removed, and items with cross loadings in two or more factors should also be eliminated (Field, 2009).

Looking at the factor analysis of those items which theoretically referred to cognitive strategies (Elaboration, Rehearsal, and Organization) the KMO value of .851 suggests that it is reasonable to run a factor analysis on the data. Bartlett’s test of sphericity is statistically significant ($\chi^2$ (210) = 5268.23, p < .001) which shows that item correlations are good enough for conducting a factor analysis. The examination of the eigenvalues as well as the scree plot show that three factors can be produced which is in accordance to the theoretical model. Results of the three-factor principal component analysis with varimax rotation are shown in Table 1.
Table 1

Results of the Rotated Component Matrix with Items Measuring Cognitive Strategies (N = 1083)

<table>
<thead>
<tr>
<th></th>
<th>factor 1</th>
<th>factor 2</th>
<th>factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>.45</td>
<td>-.04</td>
<td>.14</td>
</tr>
<tr>
<td>E2</td>
<td>.53</td>
<td>-.12</td>
<td>-.01</td>
</tr>
<tr>
<td>E3</td>
<td>.63</td>
<td>.18</td>
<td>.08</td>
</tr>
<tr>
<td>E4</td>
<td>.62</td>
<td>.16</td>
<td>.18</td>
</tr>
<tr>
<td>E5</td>
<td>.63</td>
<td>.09</td>
<td>.16</td>
</tr>
<tr>
<td>E6</td>
<td>.60</td>
<td>.03</td>
<td>.17</td>
</tr>
<tr>
<td>E7</td>
<td>.51</td>
<td>.15</td>
<td>.10</td>
</tr>
<tr>
<td>E8</td>
<td>.63</td>
<td>.06</td>
<td>.09</td>
</tr>
<tr>
<td>W1</td>
<td>.35</td>
<td>.12</td>
<td>.31</td>
</tr>
<tr>
<td>W2</td>
<td>.06</td>
<td>.18</td>
<td>.67</td>
</tr>
<tr>
<td>W3</td>
<td>.15</td>
<td>.11</td>
<td>.74</td>
</tr>
<tr>
<td>W4</td>
<td>.07</td>
<td>.12</td>
<td>.72</td>
</tr>
<tr>
<td>W5</td>
<td>.29</td>
<td>.05</td>
<td>.53</td>
</tr>
<tr>
<td>W6</td>
<td>.40</td>
<td>.02</td>
<td>.50</td>
</tr>
<tr>
<td>W7</td>
<td>.16</td>
<td>.04</td>
<td>.77</td>
</tr>
<tr>
<td>O1</td>
<td>-.01</td>
<td>.69</td>
<td>.12</td>
</tr>
<tr>
<td>O2</td>
<td>.12</td>
<td>.62</td>
<td>.09</td>
</tr>
<tr>
<td>O3</td>
<td>.10</td>
<td>.69</td>
<td>.02</td>
</tr>
<tr>
<td>O4</td>
<td>.09</td>
<td>.58</td>
<td>.10</td>
</tr>
<tr>
<td>O5</td>
<td>.09</td>
<td>.69</td>
<td>.06</td>
</tr>
<tr>
<td>O6</td>
<td>-.02</td>
<td>.76</td>
<td>.11</td>
</tr>
</tbody>
</table>

Eigenvalues

4.90  2.41  1.66

Explained Variance

14.8%  13.9%  13.9%

Cronbach’s α

.75  .77  .79

Note. Extraction method: Principal Component Analysis. Rotation method: Varimax with Kaiser Normalization. Sufficient factor loadings over the criteria .45 are written in bold.

This factor solution accounts for 42.73% of the total variance, whereas factor 1 (Elaboration) explained 14.8%, factor 2 (Organization) explained 13.9% and factor 3 (Rehearsal) explained 13.9% of the total variance. The factor analysis converged in 5 iterations. Except for the item W1 (“I have tried to learn by heart the things that could be important.”) all of the other items could be retained. The 20 items were classified based on their factor loadings as following: Elaboration (8 items), Organization (6 items), and Rehearsal (6 items). To sum up, this factor solution with three extracted factors is in line with the categories of the theoretical assumptions underlying the questionnaire construction.
Table 2
Results of the Rotated Component Matrix with Items Measuring Metacognitive Strategies (N = 1083)

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>.22</td>
<td>.46</td>
</tr>
<tr>
<td>P2</td>
<td>.08</td>
<td>.59</td>
</tr>
<tr>
<td>P3</td>
<td>.49</td>
<td>.24</td>
</tr>
<tr>
<td>P4</td>
<td>-.10</td>
<td>.66</td>
</tr>
<tr>
<td>P5</td>
<td>.40</td>
<td>.42</td>
</tr>
<tr>
<td>P6</td>
<td>.15</td>
<td>.61</td>
</tr>
<tr>
<td>P7</td>
<td>.08</td>
<td>.59</td>
</tr>
<tr>
<td>U1</td>
<td>.46</td>
<td>.12</td>
</tr>
<tr>
<td>U2</td>
<td>.56</td>
<td>.10</td>
</tr>
<tr>
<td>U3</td>
<td>.61</td>
<td>.10</td>
</tr>
<tr>
<td>U4</td>
<td>.62</td>
<td>.05</td>
</tr>
<tr>
<td>U5</td>
<td>.58</td>
<td>.22</td>
</tr>
<tr>
<td>U6</td>
<td>.32</td>
<td>.34</td>
</tr>
<tr>
<td>R1</td>
<td>.52</td>
<td>-.04</td>
</tr>
<tr>
<td>R2</td>
<td>.61</td>
<td>-.02</td>
</tr>
<tr>
<td>R3</td>
<td>.59</td>
<td>.02</td>
</tr>
<tr>
<td>R4</td>
<td>.56</td>
<td>.13</td>
</tr>
<tr>
<td>R5</td>
<td>.20</td>
<td>.05</td>
</tr>
<tr>
<td>R6</td>
<td>.05</td>
<td>.45</td>
</tr>
<tr>
<td>B1</td>
<td>.48</td>
<td>.33</td>
</tr>
<tr>
<td>B2</td>
<td>.59</td>
<td>.21</td>
</tr>
<tr>
<td>B3</td>
<td>.55</td>
<td>.27</td>
</tr>
<tr>
<td>B4</td>
<td>.51</td>
<td>.35</td>
</tr>
<tr>
<td>B5</td>
<td>.54</td>
<td>.36</td>
</tr>
<tr>
<td>B6</td>
<td>.43</td>
<td>.39</td>
</tr>
</tbody>
</table>

Eigenvalues | 6.38 | 1.80 |
Explained Variance | 20.7% | 12.0% |
Cronbach’s α | .85 | .64 |

Note. Extraction method: Principal Component Analysis. Rotation method: Varimax with Kaiser Normalization. Sufficient factor loadings over the criteria .45 are written in bold.

The factor analysis of the items theoretically referring to metacognitive strategies (Planning, Monitoring, Regulation, and Evaluation) is shown in table 2. For the factor extraction and analysis the same statistical methods and criteria were used as mentioned above. Results of the KMO and Bartlett’s test support the possibility of conducting a factor analysis. The KMO with a value of .923 suggest that running a factor analysis on these data is adequate. Bartlett’s test of sphericity also indicates good values because of its statistical significance ($\chi^2 (300) = 5733.04, p < .001$). Looking at the eigenvalues of the items as well as on the scree plot a two factorial solution is sustained. In contrast to the model-based assumptions which imply a segregation of the four aspects Planning, Monitoring, Regulation, and Evaluation, results of the factor extraction suggest (only) a two-dimensional structure of metacognitive strategies.

Factor 1 consists of items which are theoretically connected to the metacognitive strategies Monitoring, Regulation, and Evaluation except
for the item R6 ("I hurried to get everything done.") which is theoretically associated to the metacognitive process of Regulation but loaded on the second factor. This newly emerged factor contains items dealing with monitoring processes and is therefore labeled as Monitoring. On the second factor items loaded which are related to the metacognitive process of Planning except for the item P3 ("I have wondered how best to deal with the text.") which loaded on the first factor. The two-factorial solution accounts for 32.72% of the total variance. Factor 1 (Monitoring) explained 20.7% of the variance, factor 2 (Planning) explained 12%. This factor solution converged in 3 iterations.

Reliability Analysis

Based on the results of the exploratory factor analysis the items were selected to test for reliability (internal consistency by Cronbach’s $\alpha$). An item was excluded from reliability analysis if it had a factor loading less than .45 and if it had communalities less than .30. The reliability of the subscales should be $\alpha \geq .60$ to construct a reliable subscale.

**Elaboration.** The subscale Elaboration has a reliability coefficient of $\alpha = .75$. The alpha if one of the items deleted does not suggest that deleting any existing items would increase the internal consistency so all items referring to the subscale Elaboration were retained. For this reason the subscale consists of 8 items.

**Organization.** The alpha coefficient of the subscale Organization shows also a reasonable value of $\alpha = .77$. All of the six items are included in the reliability analysis and had sufficient communalities. As was the case for the first factor alpha maximization would not increase the alpha so the subscale Organization consists of six items.

**Rehearsal.** The reliability analysis of the subscale Rehearsal is conducted with all theoretical hypothesized items except the item W1 which had an inadequate factor loading. The reliability with a value of $\alpha = .79$ suggests a good internal consistency of the subscale. As before all of the six items are included in building the subscale Rehearsal.

**Monitoring.** The new emerged subscale Monitoring consists of 14 items. The items U6 ("I have taken care to be ready on time."), R5 ("If I didn’t understand a word I asked someone about its meaning."), and B6 ("I have asked myself what I would do differently the next time I had to deal with a text.") were excluded for these analyses because of inadequate factor loadings. Results of the reliability analysis show a good internal consistency with $\alpha = .85$. All 14 items were included in building the subscale Monitoring.

**Planning.** The subscale Planning contains the five items referring to the planning process based upon the model except for the item P5
(“Before starting I wondered how best to divide the work.”) with less factor loading than needed. Calculation of the internal consistency reveals a sufficient reliability with $\alpha=.64$. The alpha if deleted does not suggest that deleting an item would increase the internal consistency so all items were retained.

According to the results of the factor and of the reliability analysis one can conclude that all theoretical and model-based assumptions about the multidimensional structure of cognitive strategies proved to be reliable. The analysis of the data provides the three strategies Elaboration, Organization, and Rehearsal to be categorized as independent dimensions during cognitive strategy processing. In comparison to that, the factor analysis and tests of reliable subscales indicate that metacognitive strategies do not empirically emerge as theoretically supposed. Whereas the four strategies Planning, Monitoring, Regulation, and Evaluation were assumed to be independent dimensions of the metacognitive process the statistical analysis only reveals a two factorial structure. The metacognitive processes Monitoring, Regulation, and Evaluation cannot be empirically distinguished based on this data. So far, only the process of planning can be regarded as an independent and reliable dimension of metacognitive strategies in self-regulated learning. At this point the theoretically assumed four-dimensionality of metacognitive strategies can be questioned.

**Method Study 2**

**Participants**

Subjects of the next study were $N=1067$ students from a German Gymnasium, a high-achieving secondary school. 532 children were from fifth grade and 535 from sixth grade. There were 564 girls (52.9%) and 496 boys (46.5%), (7 missing). The age group ranged from nine to fourteen years, and the average age was 11.1 years ($SD=.80$).

**Instrument and Procedures**

The same questionnaire with the same non-fictional text and questions as in study 1 was used. It consists of 46 Items, and the seven scales Elaboration, Rehearsal, Organization, Planning, Monitoring, Regulation, and Evaluation.

**Research Question**

With regard to the results of study 1 with elementary school children and the model-based assumptions of the multidimensionality of cognitive and in particular in metacognitive strategies this study focuses also on the underlying structure of self-regulated learning. In this case we examined if the findings of the first study could be replicated with older children or if parts of especially the metacognitive processing of self-regulated learning becomes more differentiated in older children. In line
with research on developmental aspects in this area it might be possible that self-regulated learning of children between the ages of 12 to 14 years does not only show a quantitative increase in the use of metacognitive strategies but also a qualitative change. In this case we hypothesized that the metacognitive structure gets more differentiated with age. For this reason we intended to test children of the fifth and sixth grades with the same instrument to get an insight into the multidimensional structure of cognitive and metacognitive strategies in this age group.

Results Study 2

Construct Validation: Factor Analysis

For answering the research question we used the same empirical and statistical tests as well as the criteria for all parts of the analysis as in study 1. We analyzed the data of fifth graders and sixth graders separately. So we conducted two exploratory factor analyses but in the following tables and descriptions they will be reported together in order to compare the results.

Results of the KMO and Bartlett’s test of sphericity provide a factor analysis on both datasets. The KMO of the data of fifth graders with a value of .823 and with a value of .774 of the data of the sixth graders suggest both that it was reasonable to run a factor analysis on these data. Bartlett’s test of sphericity was statistically significant on the dataset of fifth graders ($\chi^2 (210) = 3100.89$, $p < .000$) and on the dataset of sixth graders ($\chi^2 (210) = 3257.82$, $p < .000$).

Looking at the factor extraction of the items referring to the cognitive strategies (Elaboration, Rehearsal, and Organization) the eigenvalues as well as the scree plot indicate a three factor solution which follows the theoretical model and the results of the first study. Results of the three-factor principal component analysis with varimax rotation were shown in Table 3 contrasting the data of children in fifth and sixth class.

The factor solution of the fifth graders accounts for 42.9% of the total variance with Factor 1 (Elaboration) explaining 16.5%, factor 2 (Rehearsal) explaining 14%, and factor 3 (Organization) explaining 12.4% of the total variance. The factor structure of the 6th graders also provides three factors with nearly the same values of explained variance (43.5%). Factor 1 (Elaboration) explained 16.5%, factor 2 (Rehearsal) explained 14.8 %, and factor 3 (Organization) explained 12.2% of the total variance. Both factor analyses converged in 5 iterations. As in study 1 all of the other items could be retained except the item W1 (“I have tried to learn by heart the things that could be important.”) which has an inadequate factor loading. The 20 items were classified based on their factor loadings and following the theoretical assumptions and the findings of the first study: Elaboration (8 items), Organization (6 items), and Rehearsal (6 items). In conclusion, these results, the factor extraction and analysis, provide the three dimensional structure of cognitive strategies. This is corresponding to the findings of the first study with elementary school children.
Table 3

Results of the Rotated Component Matrix with Items Measuring Cognitive Strategies in Fifth Graders (N = 532) and Sixth Graders (N = 535)

<table>
<thead>
<tr>
<th>Item</th>
<th>5th class</th>
<th>6th class</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>.51</td>
<td>.48</td>
</tr>
<tr>
<td>E2</td>
<td>.47</td>
<td>.57</td>
</tr>
<tr>
<td>E3</td>
<td>.56</td>
<td>.61</td>
</tr>
<tr>
<td>E4</td>
<td>.53</td>
<td>.55</td>
</tr>
<tr>
<td>E5</td>
<td>.72</td>
<td>.65</td>
</tr>
<tr>
<td>E6</td>
<td>.71</td>
<td>.63</td>
</tr>
<tr>
<td>E7</td>
<td>.70</td>
<td>.65</td>
</tr>
<tr>
<td>E8</td>
<td>.60</td>
<td>.63</td>
</tr>
<tr>
<td>W1</td>
<td>.25</td>
<td>.27</td>
</tr>
<tr>
<td>W2</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>W3</td>
<td>.19</td>
<td>.21</td>
</tr>
<tr>
<td>W4</td>
<td>.14</td>
<td>.03</td>
</tr>
<tr>
<td>W5</td>
<td>.36</td>
<td>.36</td>
</tr>
<tr>
<td>W6</td>
<td>.44</td>
<td>.39</td>
</tr>
<tr>
<td>W7</td>
<td>.17</td>
<td>.21</td>
</tr>
<tr>
<td>O1</td>
<td>.13</td>
<td>.27</td>
</tr>
<tr>
<td>O2</td>
<td>-.05</td>
<td>.00</td>
</tr>
<tr>
<td>O3</td>
<td>-.08</td>
<td>-.11</td>
</tr>
<tr>
<td>O4</td>
<td>.11</td>
<td>.18</td>
</tr>
<tr>
<td>O5</td>
<td>.08</td>
<td>-.01</td>
</tr>
<tr>
<td>O6</td>
<td>.10</td>
<td>-.12</td>
</tr>
</tbody>
</table>

Eigenvalues | 5.02 | 2.34 | 1.64 | 4.75 | 2.56 | 1.82 |
Explained Variance | 16.5% | 14.0% | 12.4% | 16.5% | 14.8% | 12.2% |
Cronbach’s α | .79 | .80 | .71 | .78 | .80 | .70 |

Note. Extraction method: Principal Component Analysis. Rotation method: Varimax with Kaiser Normalization. Sufficient factor loadings over the criteria .45 are written in bold.

Opposite to this are the findings of the factor analysis with items measuring the metacognitive strategies Planning, Monitoring, Regulation, and Evaluation. Testing the KMO and Bartlett’s test the values provide conducting a factor analysis of both datasets. The KMO of the data with fifth grade students is .888, the KMO value of the data with sixth graders is .897. Both suggest that running a factor analysis on these datasets is adequate. Bartlett’s test of sphericity also indicates good values because of its statistical significance in the dataset of fifth class students ($\chi^2 (300) = 3776.84, p < .000$) and of sixth class students ($\chi^2 (300) = 3690.01, p < .000$).

Exploring the dimensional structure of the items referring to metacognitive strategies the factor extraction with the criteria of eigenvalues over one as well as the scree plot sustain a two factorial solution. As shown in the study before the four dimensions of metacognitive processing Planning, Monitoring, Regulation, and Evaluation could not be replicated with these two datasets. Both factorial analyses suggest only a two dimensional structure of metacognition. The two-factorial solution of both datasets is shown in table 4.
In line with the results of the first study, items of the three strategies Monitoring, Regulation, and Evaluation loaded on Factor 1. This factor (Monitoring) consists of all items theoretically referring to those strategies except for the items R6 (“I hurried to get everything done.”) and U3 (“To make sure that I had understood everything I asked myself questions about the text.”) which loaded on the second factor. On the second factor all the items of the metacognitive process Planning loaded except for the item P3 (“I have wondered how best to deal with the text.”) as in the study mentioned above. The two factorial solution accounts for 35.9% of the total variance in the dataset with fifth graders with factor 1 (Monitoring) explaining 22.6% and factor 2 (Planning) explaining 13.3%. In the data set with sixth graders the totally explained variance is 36.41% with Factor 1 (Monitoring) explaining 23.5% of the variance, and factor 2 (Planning) explaining 12.9%. Both factor solutions converge in 3 iterations.

Table 4

Results of the Rotated Component Matrix with Items Measuring Metacognitive Strategies in Fifth Graders (N = 532) and Sixth Graders (N = 535)

<table>
<thead>
<tr>
<th></th>
<th>5th class</th>
<th></th>
<th>6th class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>P1</td>
<td>.16</td>
<td>.51</td>
<td>.15</td>
</tr>
<tr>
<td>P2</td>
<td>.03</td>
<td>.60</td>
<td>-.07</td>
</tr>
<tr>
<td>P3</td>
<td>.48</td>
<td>.22</td>
<td>.55</td>
</tr>
<tr>
<td>P4</td>
<td>-.15</td>
<td>.73</td>
<td>-.03</td>
</tr>
<tr>
<td>P5</td>
<td>.41</td>
<td>.39</td>
<td>.36</td>
</tr>
<tr>
<td>P6</td>
<td>.29</td>
<td>.60</td>
<td>.27</td>
</tr>
<tr>
<td>P7</td>
<td>.09</td>
<td>.64</td>
<td>.05</td>
</tr>
<tr>
<td>U1</td>
<td>.60</td>
<td>.09</td>
<td>.66</td>
</tr>
<tr>
<td>U3</td>
<td>.22</td>
<td>.46</td>
<td>.06</td>
</tr>
<tr>
<td>U4</td>
<td>.61</td>
<td>.02</td>
<td>.65</td>
</tr>
<tr>
<td>U5</td>
<td>.59</td>
<td>.06</td>
<td>.57</td>
</tr>
<tr>
<td>U6</td>
<td>.59</td>
<td>-.01</td>
<td>.64</td>
</tr>
<tr>
<td>U7</td>
<td>.74</td>
<td>.06</td>
<td>.68</td>
</tr>
<tr>
<td>R1</td>
<td>.48</td>
<td>.14</td>
<td>.58</td>
</tr>
<tr>
<td>R2</td>
<td>.55</td>
<td>.14</td>
<td>.69</td>
</tr>
<tr>
<td>R3</td>
<td>.62</td>
<td>.07</td>
<td>.68</td>
</tr>
<tr>
<td>R4</td>
<td>.55</td>
<td>.18</td>
<td>.59</td>
</tr>
<tr>
<td>R5</td>
<td>.10</td>
<td>.14</td>
<td>.09</td>
</tr>
<tr>
<td>R6</td>
<td>.04</td>
<td>.55</td>
<td>-.03</td>
</tr>
<tr>
<td>B1</td>
<td>.53</td>
<td>.35</td>
<td>.48</td>
</tr>
<tr>
<td>B2</td>
<td>.63</td>
<td>.03</td>
<td>.69</td>
</tr>
<tr>
<td>B3</td>
<td>.53</td>
<td>.26</td>
<td>.54</td>
</tr>
<tr>
<td>B4</td>
<td>.55</td>
<td>.34</td>
<td>.52</td>
</tr>
<tr>
<td>B5</td>
<td>.54</td>
<td>.40</td>
<td>.51</td>
</tr>
<tr>
<td>B6</td>
<td>.54</td>
<td>.34</td>
<td>.37</td>
</tr>
</tbody>
</table>

Eigenvalues | 6.77 | 2.20 | 6.61 | 2.49 |
Explained Variance | 22.6% | 13.3% | 23.5% | 12.9% |
Cronbach’s α | .88 | .72 | .89 | .73 |

Note. Extraction method: Principal Component Analysis. Rotation method: Varimax with Kaiser Normalization. Sufficient factor loadings over the criteria .45 are written in bold.
Reliability Analysis

As in study one the items are selected to test the reliability based on the results of the exploratory factor analysis and their factor loadings. All analyses showed a good internal consistency for all subscales as follows. The alphas if one of the items deleted does not suggest that deleting any existing items would increase the internal consistency so all items loading on one subscale were retained.

Elaboration. The subscale Elaboration has a reliability coefficient of α=.79 in the data of fifth and an alpha of .78 in the data of sixth graders. The subscale consists of 8 items in both datasets.

Rehearsal. The reliability analysis of this subscale was conducted with all theoretical hypothesized items except the item W1 which had an inadequate factor loading as in study 1. The reliability with a value of α=.80 in both analysis suggest a good internal consistency of the subscale. As before all of the six items are included in building the subscale Rehearsal.

Organization. The alpha coefficient of the subscale Organization shows also a reasonable value of α=.71 in the fifth class data set and α=.70 in the sixth class dataset. All of the six items are included in the reliability analysis and had sufficient communalities.

Monitoring. The newly emerged subscale Monitoring consists of 16 items in the data of the fifth graders. The items U6 (“I have taken care to be ready on time.”) and R5 (“If I didn't understand a word I asked someone about its meaning.) are excluded from analyses because of inadequate factor loadings. Results of the reliability analysis show a good internal consistency with α=.88. In contrast to this the analysis of the data measuring Monitoring in sixth graders only 15 items loaded sufficiently on this factor. The item B6 (“I have asked myself what I would change the next time when I will read a text.”) has an inadequate factor loading and was therefore excluded. The internal consistency of this subscale is α=.89.

Planning. Looking at the data of the fifth graders, the subscale Planning contains the five items referring to the planning process based upon the model except the item P5 (“Before starting I wondered how best to divide the work.”) with less factor loading than needed. Calculation of the internal consistency reveals a good reliability with α=.72. Analyzing the data of the sixth graders the item P5 has a sufficient factor loading on the factor. In this case the subscale Planning consists of six items and also had a good internal consistency with a value of α=.73.

To sum up it can be concluded that in line with theoretical models of self-regulated learning the three-dimensional structure of cognitive strategies was confirmed by the results of these analyses comparing data
of fifth and sixth grade students. In line with the theoretical assumptions we replicated the three strategies Elaboration, Organization, and Rehearsal as independent and reliable dimensions of the cognitive learning strategies. In contrast to this, the findings regarding the dimensionality of metacognition only suggest a two-factorial structure. Whereas several models of self-regulation assume a three or four phase structure of this metacognitive aspect the empirical data provides only evidence for monitoring and planning processes - at least for children between nine to fourteen years. According to this finding we had no hints for a more differentiated metacognitive processing than in elementary school children. The results hardly provide the two-dimensional structure of metacognitive strategies in school children.

**Discussion**

Our investigations clearly show that the assumed model of three separate factors for cognitive learning strategies can be found - like in many other studies - without doubt also in our data. In this respect it replicates the findings already reported by Weinstein and Mayer (1986), Wild and Schiefele (1994) and Garcia and Pintrich (1996).

On the other hand, the assumed model of four factors for metacognitive learning strategies was neither found in the first nor in the second study. What could be possible and plausible explanations for these results? We suggest that neither the formulation of our items nor the age of the sample is responsible for the misfit of theoretical model and empirical data with regard to the metacognitive strategies.

Our results might resemble to the model of Weinstein and Mayer (1986) which does not separate metacognitive strategies into further subgroups. But our data reveals that Planning as a metacognitive strategy can be separated from other metacognitive processes. Several other studies also had difficulties to separate metacognitive strategies (e.g. Garcia & Pintrich, 1996; Wild & Schiefele, 1994). Whereas for the cognitive subscales three different dimensions are clearly empirically separable, metacognitive processes in self-regulated learning are more interwoven and highly correlated. Pintrich (2000) states that there are four different phases in which different metacognitive strategies play their part. According to Pintrich (2000), these phases are not necessarily independent and distinct. Our analyses reveal at least one separable phase which is Planning. The three other theoretically assumed metacognitive strategies are highly correlated (e.g. monitoring and regulation) and interwoven in the learning process and can not be separated empirically from each other (as already Garcia and Pintrich (1996), Spoerer (2004) and Wild and Schiefele (1994) have reported). Therefore it has to be questioned if the underlying theoretical model should be changed. In our opinion changes in the theoretical models are not necessary.

On the contrary: The assumption of a highly differentiated process of metacognitive regulation helps to explain what is useful and necessary
when trainings of metacognitive regulation are planned and administered. The process of metacognitive regulation can be illustrated easily by the different facets in a complex process, but like in a complex dance a separate step alone is not a representation for the nicely coordinated process. Therefore we would recommend not to reduce the theoretical model but to make clear that the underlying mechanisms of metacognitive regulation are rather helpful to explain and illustrate a complex picture of Regulation! But it should be made clear as well, that one strategy (e.g. regulation) is senseless without another strategy (e.g. monitoring).

To sum up: Our studies are in line with the theoretical models of self-regulated learning of Weinstein and Mayer (1986) and Garcia and Pintrich (1996). Both models postulate three different dimensions of cognitive learning (rehearsal, elaboration and organization), which we could clearly separate in two studies with children aged eight to fourteen years old. With respect to metacognitive processes the models differ. Whereas Weinstein and Mayer (1986) do not postulate any sub-dimensions of metacognitive learning, Pintrich (2000) conceptualize four different but interwoven metacognitive phases. Our data shows that only Planning is separable from the other metacognitive phases postulated by Pintrich (2000). However, we recommend to adhere to the model of Pintrich (2000) in order to illustrate different (but interwoven) processes of metacognitive self-regulation in training situations.

References


**About the Authors**

Stephan Wernke is a doctoral candidate at the University of Oldenburg in Germany. His research interests are self-regulated learning, learning strategies, metacognition, and self-concept.
stephan.wernke@uni-oldenburg.de

Dr. Uta Wagener is research assistant at the University of Osnabrueck in Germany. In her research she is focusing on self-regulated learning in young children, learning strategies, metacognition especially in observational studies.
uta.wagener@uos.de

Andrea Anschuetz is doctoral student at the University of Oldenburg in Germany. Her research interests are epistemic beliefs, learning strategies, attributional theories and its training.
andrea.anschuetz@uni-oldenburg.de

Prof. Dr. Barbara Moschner is professor of Learning and Instruction at the University of Oldenburg in Germany. Her areas of interests are self-regulated learning, self-concept, epistemic beliefs, attributional theories, and the research of mistakes in school settings.
barbara.moschner@uni-oldenburg.de
The Types and Nature of Questions vis-à-vis Students’ Test-Taking Skills as Significant Indicators of Second Language Examinees’ Performance on the TOEFL-ITP Reading Comprehension Sub-Test

Analiza Perez-Amurao
Mahidol University International College
Salaya, Thailand

Abstract
This study examines the reading performance of selected students at the Pre-College program of the Mahidol University International College (PC-MUIC) as they are required to attain a score of 520 in the TOEFL-ITP (or equivalent performance in IELTS) to enter MUIC. Specifically, this research aims to evaluate whether the reading skills that examinees possess correlate with successful performance on the Reading Comprehension sub-test of the TOEFL-ITP. Only TOEFL-ITP Reading Comprehension Sub-test performance has been considered in this study as IELTS is not taught or administered in the Pre-College program. This study makes use of descriptive qualitative-quantitative design relying heavily on the following instruments for data collection: Commercial-based test-prep texts (Reading Comprehension Sub-section), Schraw and Roedel’s Levels of Difficulty (1994), the researcher’s modification of said band, the respondents’ scores per question type, tabulations of the respondents’ scores based on the levels of difficulty of the items and the question types used in the test, focused interviews with the respondents, and retrospective journal entries of the researcher. This study aims to shed light on issues surrounding how second language learners’ reading skills affect performance on standardized tests such as TOEFL. This study specifically seeks to provide MUIC PC instructors empirical data that would help them understand their own students’ reading difficulties which, consequentially, will aid them address teaching-learning issues.

Key words: TOEFL, reading skills, test-taking

Introduction

In an EFL/ESL classroom, cases of a reading teacher facing the enormous and challenging task of making students love reading as the students themselves hesitate or refuse to embrace the benefits of the exercise are no longer new. Many times, a reading teacher finds himself frustrated at the end of the day as he finds it hard to make the students to even open the pages of their reading material (Snow, Griffin, & Burns, 2005). This, of course, is complicated by the fact that the teacher has no choice but to push the students even further not only because they need to learn and use a number of reading strategies at the end of the term, but also because they need to acquire reading skills and test-taking strategies for them to achieve satisfactory scores in the assessment stage.

In situations where the students’ performance on high-stakes tests both culminates in one academic exercise, on one hand, and commences in another, on the other, both the learner and the teacher come to a point where obtaining the required score becomes the chief objective. In circumstances like this, the challenge of realizing the chief objective lies
heavily on the shoulders of the teacher. This, of course, does not mean that the students are freed of their responsibility. This simply means that apart from the pro-active attitude expected of them being the test-takers and direct recipients of said pedagogic exercise, the academic community also looks forward to seeing the teacher take a more hands-on approach if only to moderate the task and eventually accomplish the goal (Dole, 2004).

In an effort to assist reading teachers, academicians and theorists who are monitoring this field for future academic discourse, this study was conducted to investigate and consequently provide corroborated data on learners’ reading and test-taking skills as revealed by their scores in the Reading Subtest of their complete practice test. This paper conceptually aims to validate the importance of drawing on appropriate reading strategy albeit indirectly carried out in this study as correlated to achieving satisfactory scores in the said standardized test within the Reading Comprehension context.

This paper aims to establish a premise underscoring how the subjects’ test performance directs educators to seek appropriate measures namely, but not limited to, the following diagnostic objectives: to identify the types of TOEFL-ITP Reading Comprehension questions examinees failed to understand and respond to accurately; to determine the nature of the types of questions examinees performed unsatisfactorily on; to identify the test-taking strategies examinees failed to use before and during the testing period, and; to identify the strategies that they presumably used successfully.

Likewise, this study underpins the need to attain the following pedagogic objectives: to construct effective activities based on the students’ test-taking strategies that were judged to be successful (Focus on activities); to provide students with a cognitive framework for improving their test performance (Focus on self-development), and; to develop a test-taking strategies syllabus based on the diagnostic and pedagogical data that has been identified by this study (Focus on teaching materials).

The growing demand for the English language has resulted in an increase in the use of high-stakes international tests such as TOEFL, IELTS, TOEIC and the like. While more and more users of the English language are relying heavily on these tests for various reasons, including for admission to universities and for employment abroad, Andrew Cohen (2006) said that for the past 30 years, the focus of research in second language (L2) assessment was primarily on testing results. There was little attention given to what the test-takers do to arrive at the right answers and how the assessment matched the skills it was supposed to test. While there have been some studies conducted focusing on test-taking strategies, results of this researcher’s inquiries revealed dearth in studies within the Thai context on L2 assessment looking into the relationship between the examinees’ language skills, their test-taking strategies and the effectiveness of strategy instruction so as to aid test-takers’ performance on high stakes standardized tests.

In a study of a Thai university students’ performance on TOEFL, Marukatat and Bunnag (2003) found out that based on the scores
gathered from nine of the 10 Asean member-states whose TOEFL results were studied, the Thai students performed second from the last. The outcome showed how the Thai students were surpassed by their peers from the neighboring countries particularly from Burma, Vietnam and Cambodia. Various reasons were cited. In some other studies within the Thai context, one of the strongest reasons that accounted for such a poor performance was the country’s lack of a strong reading culture, even in its native tongue. “It is unlikely that students have ‘reading’ models, a factor that may have a significant impact on the students if the teachers themselves are not seen to be readers” (Eskey & Grabe in Wisaijorn, 2008). Another reason considered very influential is that since many traditional Thai classrooms are teacher-controlled, the students have not been trained to think independently for themselves. Because of excessive emphasis on social conformity, i.e., “submissive students who do not ask questions are seen as well-behaved,” the students, in turn, have become very passive, afraid of asking even the most relevant questions about their readings. Students who are “critical and analytical and who reason with their teachers are often viewed as aggressive and disobedient, and have trouble fitting into the Thai education system” (Chareonwongsak, 2007, p. 4). In situations like this, one can easily deduce how emotional factors control learners’ accomplishments and actions (Alvermann & Guthrie, 1993). Research shows that when children are able to make positive associations with reading, they tend to regard the task as something rewarding and enjoyable. However, when they see negative connections between reading and what actually happens inside the classroom, “their achievement tends to suffer. These children will either avoid reading altogether or read with little involvement” (Henk & Melnick, 1995, p. 470).

On a related note, students from Mahidol University International College (hereafter referred to as MUIC) are required to attain a score of 520 in TOEFL (or equivalent performance in IELTS) to be admitted to the university. As preliminary observation based on some learners’ past performance on TOEFL exams and review classes yielded, PC TOEFL instructors found out that learners have the most difficulty doing the Reading Comprehension Subtest. This problem is linked to the Thai students’ ability to cope with simple communication in their own area, and in their inability to use advanced language (Fredrickson in Wisaijorn, 2006). In line with this, a study on the students’ language proficiencies, test-taking strategies and strategy instruction can prove to be of immense value as it will provide the stakeholders, i.e., students, teachers, curriculum designers, practical information on how to improve the examinees’ performance. This primarily explains why this study placed a singular focus on the students’ performance on the Reading Comprehension Subtest.

Second language assessment research, for the most part, has focused on issues that deal primarily with the outcomes of language testing, item performance, test reliability and the like (Cohen, 2006). While results of these studies have served great roles especially useful for academicians, test constructors and even test-takers themselves, the fact
remains that L2 assessment has focused very little on the types and nature of questions L2 examinees find difficult responding to. Below are brief discussions of studies and results done in L2 assessment. These research results are presented with the aim of providing a bigger picture of what was already done and what is still needed to be done in said field.

Carol Fraser’s (1999) report entitled “Lexical Processing Strategy Use and Vocabulary Learning through Reading” looked into the lexical processing strategies (LPSs) used by L2 learners as they tried to make sense of new words while reading and sought to see the effect of such strategies when learning unfamiliar lexical items. This study aimed to increase one’s understanding of the role of LPS in as far as how they can help in the reassessment of the current academic practice. Results showed that generally, instruction using LPSs indicated potential for enriching one’s vocabulary. The study also demonstrated that the use of some LPSs leads to retention rates higher than the others.

In 2002, Abanomey (In Cohen, 2006) conducted a test-taking strategy research exploring some features of a test format and checking on how influential the use of authentic texts is against inauthentic ones in a reading test. The inquiry looked into whether authentic texts would impact the manner in which test-takers employ test-taking strategies. It also explored the differences, if there are, between test-takers reading authentic texts and those reading inauthentic ones in their use of “bottom-up (text-based) and top-down (knowledge-based) strategies.” Results showed that text authenticity did not affect the number of strategies used with them, although it did affect the manner in which examinees used the test-taking strategies.

In a related study in 2002, Liz Hamp-Lyons and Alan Davies carried out a research project with a hypothesis stating “that international English tests are biased: by that we mean that they systematically misrepresent the ‘true scores’ of candidates by requiring facility in a variety of English to which whole groups of candidates have not been exposed.” This study primarily focused on issues that relate to assumptions about International English (IE) and World Englishes (WEs) views. Realizing that their data set was not sufficient for them to come to any conclusions, as they were only left with one data that appeared to support their WEs hypothesis, Hamp-Lyons and Davies suggested instead that a further study be done because although the “bias” on the basis of their research was “not proven,” it could not be dismissed either.

In the context of a study done in 2002 involving the use of TOEFL, David Qian’s (2002) investigation validated the significance of breadth and depth of vocabulary knowledge in reading comprehension within the scholastic setting. The study found out that the “dimension of vocabulary depth is as important as that of vocabulary size” in forecasting test-takers’ performance on academic reading.

In a more localized research setting in Thailand, Patarareeya Wisaijorn’s (2008) “Strategy Training in the Teaching of Reading Comprehension” looked at the country’s L2 reading situation and saw how current classroom practices have been directly influenced by the Thai
educational culture. Wisaijorn found out that having the greatest tendency to be teacher-centered and teacher-directed, the Thai education system negatively influences reading in English in most Thai classrooms such that “weak performances in reading in English indicate difficulties in fulfilling the demands of their [students’] studies.”

While there have already been a considerable number of researches in L2 assessment, studies that explore on the correlation between the types and nature of questions as significant indicators of L2 examinees’ performance on the TOEFL-ITP Reading Comprehension subtest remain deficient. So much so that the context in which this study was done is in Thai, a culture that has been generally labeled in a number of local researches to have a very poor reading ethic either in its own language or in a foreign one. Situated in a milieu where the education system is highly influenced by its culture, the Thai university students’ reading skills seem to be performing unsatisfactorily. It is within this same vein why this research aims to pioneer investigations on Thai university students’ performance on a high stakes test, particularly TOEFL. This study is anchored on the premise that by knowing and understanding the types and nature of questions and test-taking strategies used by the examinees, said variables would serve as significant indicators of their performance on the Reading Comprehension component of said test and, consequently, will aid in the curriculum and materials development.

Method

Design

The study made use of the descriptive qualitative-quantitative design. Using the scores per question type, this research observed and measured the behavior of the participants in relation to the specific types of questions they performed on successfully and unsuccessfully.

Purposive sampling was used as the main basis of selecting the participants. This sampling method was used because of the very nature of the investigation which is to look into the reading skills of a specific group of students placed under a bridge program in the Thai educational setting.

Instruments

The investigation relied heavily on the following instruments for data gathering: commercial-based test-prep texts from Barron’s TOEFL Strategies, a commercial-based test-prep material published by Barron’s Educational Series, Inc. (particularly the Reading Comprehension subtest); Schraw and Roedel’s Levels of Difficulty (1994) with researcher’s modification; the subjects’ scores per question type; tabulations of the respondents’ scores (Based on the levels of difficulty of the items and the question types used in the test); focused interviews with selected students, and; retrospective journal entries.
Participants

This study involved three groups with a total of 43 students admitted to Level Four (4) of the PC Program during Quarter Four (4) of AY 2008. Level Four (4) students are those whose language skills have been identified to be in the Intermediate and Upper-Intermediate levels. Students in this group are those who are expected to join mainstream university classes at Mahidol University International College assuming that they get an average of D+, equivalent to 65%, in their writing, reading, and listening/speaking non-credit PC classes, and achieve a TOEFL score of 520. The selection of the students admitted to the program was primarily based on the scores they obtained after taking MUIC’s College Entrance Examination. Being an international college, MUIC requires that apart from passing the entrance examination, students should also obtain a TOEFL-ITP score of 550 or above. When students fail to satisfy the admission requirements, they are recommended for enrolment at the Pre-College Program (hereafter referred to as PC), a bridge program of the MUIC which aims at improving the students’ English language and mathematical skills.

Procedure

The investigation was done by, first, extracting the questions from the Reading sub-test and tabulating them based on the subjects’ correct and wrong responses. Second, using the initial tabulation, all the 50 Reading Comprehension questions were arranged according to their levels of difficulty the basis for which was Schraw and Roedel’s three levels of difficulty. In their report, Schraw and Roedel validated the use of a band to identify the levels of difficulty of test items based on the success rate of the entire research population. Success rates were determined based on how difficult all the 43 respondents found each question. In this study, an item that had 30%-50% of 43 students responding to it correctly was labeled Difficult (D) while an item that had 50%-70% of 43 students responding to it correctly was marked Moderately Difficult (MD). An item that had 70%-90% of 43 students responding to a question correctly was classified as Easy (E).

Third, the researcher grouped the scores based on the question types and amended the band by organizing the levels of difficulty into four. The fourth level was added to accommodate scores of students that did not fall under the three levels included in Schraw and Roedel's study. Grouped this way, the scores and, implicitly, the students’ reading skills were explored and analyzed. In the analysis, the examinees’ test performance was discussed in light of existing language learning-teaching issues and theories and the researcher’s general familiarity with the examinees’ knowledge and/or lack of knowledge of test-taking skills.
Data Analysis

To confirm research findings, the researcher also triangulated her data with two other research instruments, namely, focused interviews with selected students and retrospective journal entries vis-à-vis an ethnographic observation of said interviewees. This way, biases arising from interpretations based on pure assumption are done away with.

To fulfill one common-knowledge principle for testing which is to identify a student's areas of relative strength and weakness in subject areas, analyzing and making sense of what test results imply proves to be one of the most fundamental steps. In analyzing the results of the practice test taken by the subjects in this study, the researcher initially considered the use of basic statistical analysis. After tabulating the scores, however, she and her statistician decided that even a simple tabulation of scores and other related information triangulated with the use of other research instruments, such as focused interviews and retrospective journal entries, would be enough to provide her with the most salient data needed to arrive at significant statements in conjunction with the research objectives of said study. Focused interviews provided the researcher students’ insights about the types and nature of questions they found difficult. Said interviews allowed her to look further into why they found some questions difficult to tackle. Taking advantage of the privileged position in noting down every single detail one can get from an insider’s perspective, the researcher also kept journal entries to substantiate findings and interviews.

To check on the subjects' performance on the test, the researcher tabulated the scores and ranked them based on the correct and wrong responses they gave. This would allow anyone to see that the test items could be interpreted in isolation from the others despite the fact that they came in clusters. This means that even if one set of questions was based on one same passage, each test item could be interpreted independently as answers to the succeeding questions were mutually exclusive and not dependent on the previous one. Having ranked the scores, the researcher then identified the level of difficulty of each test item by using the categories used by Schraw and Roedel in their study. As discussed previously, because the categories in Schraw and Roedel’s study (1994) were only comprised of three levels of difficulty, those whose scores that fell lower than what was originally cited in the 1994 study were put under another level, the result of the researcher’s modification of the band. In so doing, scores that fell between the 10%-30% range were classified as Extremely Difficult (ED).

Results

Results show that the spread of scores was very wide illustrating that the respondents’ performance on each question type was not consistent. Because of this, the researcher deemed it more practical to offer an analysis and interpretation of results based on the correct
responses given per type of question rather than on the overall ranking of the test items irrespective of question types.

To address directly the goal of this study, the foregoing discourse is hinged on a number of diagnostic and pedagogic objectives the researcher found to be achievable. There are some diagnostic and pedagogic objectives, however, which the researcher reckoned not attainable as the data collected were not sufficient to provide substantive analysis and interpretation. A more comprehensive study on the respondents’ test-taking and learning strategies may be conducted later to attain some of these goals. As such, issues surrounding the third and the fourth diagnostic objectives are not dealt with in the succeeding discussion. On the other hand, even if the first and the third pedagogic objectives are directly related to the respondents’ test-taking and learning strategies, this study, nonetheless, offers to provide information on them which, although not necessarily directly related, is allied, at the very least.

The one last, but equally important, issue that readers of this paper need to be reminded about is the testing conditions under which the respondents of this study were asked to go through. In the actual TOEFL exam, test takers are asked to complete all the three sections of the exam in one meeting which normally takes a little over two hours. No break in-between test sections is allowed. In this study, however, in as much as the proctors/lecturers of the practice test tried to establish and maintain the same testing conditions observed during the actual TOEFL exam, the practice test was divided into two, the first two sections of which, namely, Listening and Structure and Written Expression, were taken in one regular classroom meeting. The Reading Comprehension section was taken the following day, during the regular classroom session, too. This set-up needs to be factored in because slight differences in testing conditions are normally anticipated to have effects on test-takers’ performance and test results.

The presentation of the results of this study will be two-pronged such that the question type and the levels of difficulty of each type will be provided to give a better and clearer context.

Question type 1 is on Facts & Details. Questions of this type are those that require answers that are directly stated in the passage. Ordinarily, answers to these questions may be arrived at even without having to draw a conclusion about a text read.

Under the Facts & Details test type are four questions which were found to be Extremely Difficult with the top-ranking question yielding a mere 16.28% success rate among the respondents. Consistently, this item also ranked 2nd in the overall ranking of test items.

Although the test items seemed to be knowledge questions at first, their very nature actually asked examinees to go beyond recognizing details in the passage. Apart from the tabulations demonstrating items 24, 26, 27 and 31 as Extremely Difficult ones, ethnographic observation further revealed that the reason why the respondents found them as such was that these questions required them to synthesize and discriminate information, a higher-order thinking skill the researcher conjectured the
respondents had not acquired yet at the time of the test. Apart from the lower-order thinking skills details cited in both the stem and the options, which fell under Knowledge and Comprehension (Bloom, 1956), the test takers were supposed to take note of higher-order thinking skill points that particularly required them to do an analysis and a synthesis before the correct answer for each could be arrived at. On top of this, it is also important to note that while the examinees found said four questions to be Extremely Difficult, the success rate of the students on each test item varied both within the same level of difficulty and across the other levels. This means that the variation actually stemmed from the degree of specificity of each question checking on how familiar each respondent was with the language and content of each question regardless of the students’ being homogenously grouped in the PC Program and the homogenous kind of instruction they received prior to taking the test. Question 24 came out with the least rate of success at a meager 16.28% making it more difficult by 10% compared to Questions 26 and 31. Questions 26 and 31 were of equal rank turning up 25.58% only of the 43 respondents who exhibited mastery of said skill at the time of the test. Question 27, with a success rate of 27.91%, indicated that knowledge of this question was slightly higher by 2.33%.

Questions 18, 45, 4, 6, 9 and 22 comprised another sub-category of a Facts & Details question type. These items were considered Difficult. The researcher gathered that although the questions merely asked for facts, generally speaking, they were not simple Wh-questions. She further gathered from focused interviews that additional test item descriptors also known as expanders such as “…what common characteristics distinguished the careers of the Mayo brothers?” added to the degree of specificity of the question that the respondents had to deal with as opposed to the usual Wh-question which could have gone as simple as this: “What common characteristics did the Mayo brothers have?” Following the expanded synthesis question format, Questions 18, 4, 9 and 22 appeared complex for the examinees to readily understand and answer correctly. Another test item, question 45, on the one hand, was difficult, nevertheless, because although the descriptor came out short and simple, the options required the examinees to process and eventually analyze each of them more thoroughly. The analysis required by each option made the item seem a trick question for them. Question six, on the other hand, no matter how simple and straightforward the question was, came tricky as well because the students had to be very keen on details. When asked, students said that the error may be traced back to their not having checked each distractor very carefully.

Only two Facts & Details Questions fell under another sub-category, namely the Moderately Difficult Level, with success rates of 55.81% and 69.76%. Despite having a difference of 14%, Questions 38 and 16 were both categorized as Moderately Difficult items only primarily because the degree of specificity of the descriptors did not appear to be as high as those of the questions found in the Extremely Difficult and Difficult test items. Although the nature of these questions did not vary
much from the previous ones (questions 24, 26, 27 and 31) labeled as Extremely Difficult and Difficult, the additional descriptors and expanders of these questions were simpler and fewer compared to those of the former such that the underlined words in the questions “...what purpose do the fine hairs on the body of the bee serve?“ and “... what does marketing research include?” posed to be much easier to comprehend. On this, the respondents expressed their preference for simpler and shorter questions.

Questions 39 and 37 came out to be another sub-category. They placed under the Easy Level. Although the basic aim of this question was to require the examinees to extract facts and details from the text, tabulations indicated that these two questions turned out to be Easy for them. The researcher’s observation revealed two possible two reasons, later confirmed by the respondents during the focused interviews. First, both the stem and the options—the distracters and the best answer—were stated in simple form. Second, both questions were plain knowledge questions—those that touch on the lower-order thinking skills—asking the examinees to recall information from the text, virtually the lowest form of task test takers normally do cognitively.

The second type of question has to do with making an Inference. An inferential question asks an examinee to comprehend an idea or argument that is strongly implied but not directly stated in the text.

The question type with the third highest number of items included in the test—inference questions 44, 49, 3, 21, 47 and 34—were found to be Extremely Difficult and ranked as the second most difficult ones. Unlike questions in expanded form, those which asked the test takers to integrate additional pieces of information provided in the stem of the question, questions 44, 49, 3, 21, 47 and 34 were stated following a much simpler format. Despite the simplicity of their format, however, these questions turned out to be the second most difficult questions the examinees found. This difficulty can be attributed to the fact that “inferential comprehension questions measure interpretation. These items require one to ‘read between the lines’ or even ‘beyond the lines’ combining past knowledge and familiarity with text information. These were synthesis questions, the highest form and most challenging one based on Bloom’s Taxonomy of Learning (1956).

In the case, however, of the next two groups of Inference Questions based on the levels of difficulty aptly classified as Difficult and Moderately Difficult, although considered also as Inference Questions, they did not turn out to be as complicated as the others for two interconnected reasons. First, the questions simply asked the test takers to infer information that could be easily traced to simple facts and details cited in the passage. Second, because of this, the options were not as abstract as was the case of the questions 44, 49, 3, 21, 47 and 34.

In other words, the nature of the inference questions such as items 42, 29, 33, 15, and 36 required the use of information that was more explicitly stated in the passage. It was a less difficult task required by the previously mentioned items, which was to merge concepts and produce something implied in the text.
The third type of question is about identifying the **Main Idea** of a given text. This is a type of question about the overall idea expressed in the passage. All this asks is the primary point which the writer is trying to convey.

Only three **Main Idea Questions** were included in this test. Because of the small number of questions, results particularly involving said test items cannot be considered conclusive. Nevertheless, the use of said items in this test was deemed significant if only to establish a baseline data unique to the respondents’ learning context. Of the three **Main Idea Questions** included in this test, only one item turned out to be **Extremely Difficult**, such that only 25.58% got the correct answer. This question, item 12, was a main idea question which turned out to be difficult. Apart from having to synthesize the meaning of said question, students ultimately had to make connections between the text and a related subject matter. In so doing, the test takers were actually asked cognitively to use the information they read about in another related context or situation. Interview results revealed that they were not used to making and/or establishing connections and patterns the way they were asked to do on the test. Part of the difficulty may also be attributed to the students’ having been so used to simple Main Idea Questions such that requiring them to make associations seemed too difficult a task for them to fulfill at the time. This confirms one of the findings of Wisajorn’s study (2008) concerning features of Thai education saying that “students may find it difficult to develop skills in creative thinking, independent and alternative learning, questions and/or discussion.”

Interestingly, the gap between Question 12 (**Extremely Difficult**) and Question 1 (**Difficult**), and Question 1 (**Difficult**) and Question 50 (**Moderately Difficult** ranged from approximately 13.95% to 13.96%, a consistent difference across three levels of difficulty. What can be seen as a give-away in Question One was that the format of the question was the most common the test takers normally encountered especially in their practice tests, seat works, and assignments. It did not come out as a surprise that they found the same line of questioning easy to handle. Also, the examinees found Question 50 shown above clear-cut and, therefore, easy. Both the stem and the options were precise such that matching the test item and the answer did not give the test takers a hard time.

**Organization** is the focus of the fourth type of question. This question asks a test-taker to determine how the ideas in one paragraph relate to the ideas in another paragraph. In this sub-category, Question 11 turned out to be **Difficult**. It may be assumed at first that the item, question 11, looked easy to handle, as it only required the examinees to identify the exact location of the information in question. Tabulations disclosed, however, that this seemingly easy question was actually extremely difficult for the examinees in general. Students’ reactions showed that this difficulty can be traced primarily to how the stem was crafted. While it mainly tested a test taker’s knowledge of the organization of the text, it turned out more complicated than expected because of its expanded format no matter how simple the question was. It required the
examinees to look for some pieces of information considered crucial to determining the exact location of the data in question. The students said that the test item seemed too long a string of a question. As such, the researcher gathered that “for the expansion” and “of the practice” expanded what could have been a simple one. Such a difficulty was very similar to what the test takers commented on regarding test items 18, 4, 6, 9 and 22.

If students could not identify and/or establish the relationships between and among the parts within a long string of a question, finding the correct answer to it then was simply impossible. This explains as well why this particular test item, although it virtually asked test takers to perform the same task similar to what the students did on Question 32, proved to be more difficult.

Another sub-category of a test type about Organization is the Difficult Level. Although slightly different in terms of percentages, questions 43 and 32 were found to be Difficult. The two test items below, of course, are not exactly of the same format. Question 43 straightforwardly asked the examinees to identify the possible source of the passage in question thereby requiring them to make some judgment and eliminate the distracters in favor of the best answer. Question 32, on the other hand, simply asked the test takers to map out the passage to enable them to figure out where exactly a piece of information in question could be found. Different the questions might had been, the examinees found both of them difficult. What appeared to be more interesting, however, was that Question 43 turned out to be more difficult than Question 32, the reason for which might be attributed to the nature of the former. It did not only simply make the test takers think about the plausibility of all of the options. It also required them to do the following: One, to give meaning/s to each of the options; two, to judge which of the options made the most sense, and; three, to eliminate the distracters so as to choose the best one among them. Required to do all three sub-tasks to answer a single item, the students admitted that the steps seemed too daunting a task for them to do. This result came out to be in synch with the researcher’s observation of the students’ performance on the test.

The fifth type of question is on Referents. Typically, this type of question asks an examinee to determine which noun a pronoun refers to and/or vice versa. Although of different percentages, Questions 23 and 20 both emerged as Easy items. This could be attributed to two main reasons: one, the students have always had this type of question in their reading quizzes hence their level of familiarity with it was high, and, two, this type of question fell on the knowledge category of questions in reference to Bloom’s Taxonomy. The use of knowledge questions in this part of the test, items that checked on and activated the students’ lower-thinking skills, required students to make a simple matching of the term in question and its referent.

The sixth question type asks about Tone & Purpose. A question about the ‘tone’ requires the examinee to identify the emotion the writer is
trying to show in the text whereas a question about the ‘purpose’ asks what the author is trying to do in the passage.

Although what can be said about this type of question could not be taken conclusively as it was the only item of this kind in this test, giving it a serious thought may greatly help future studies. It ranked, however one puts it, as the 3rd most difficult question making it an Extremely Difficult item with only 20.93% success rate. Similar to how the examinees performed on items that tested their ability to make good inferences, which were classified either as Difficult or Moderately Difficult, this particular question that checked on the students’ ability to sense the Tone and Purpose of the text and/or author proved to be one item they needed to put more attention to. This runs consonant with the researcher’s observation and interview results that because the respondents were not in the habit of reading texts in English—aside from what was given them within the classroom setting—they had very limited background knowledge. Having rich background knowledge, on the other hand, could have been very helpful in their understanding of the author’s intention in the text they read. Schema theory states that “reading comprehension is an interactive process between the text and the reader’s prior background knowledge” (Adams and Collins, 1979; Rumelhart, 1980). This leads to an understanding that their schema at the time of the test was not sufficient to help them understand better the texts they read.

What seemed to have compounded the situation was the fact that the test-takers not only had to sense the tone and purpose behind the text. They also had to provide a reason, a task that required them to approach the item employing their cognitive skills. In so doing, they needed to give meaning to each of the options requiring them to synthesize each before eliminating the distracters. To add to that, the stem itself of question 19, which used highly specific terms such as “legally binding agreements”, if not comprehensible and familiar enough to the test-takers could have been a major source of difficulty.

The last test type is on Vocabulary. These items basically ask for word meanings doing which can be done in a number of ways, namely, using structural clues, understanding meanings from word parts, and finding definitions from context clues.

Based on at least three consecutive actual TOEFL exams administered in the past, the PC TOEFL lecturers observed that Vocabulary questions, followed by Facts and Details, had the most number of questions in the test. It is actually a trend in practice tests that holds true even in the actual TOEFL-ITP exam. This explains partly why Vocabulary questions also had the most number of questions in the Extremely Difficult category. Of greatest importance, however, in the discussion of the examinees’ performance on this test are the following: (1) the word “consistent” turned out to be the most difficult word yielding a 13.95% success rate, and: (2) eight vocabulary items out of 21 questions occupied the Extremely Difficult category.

Based on their levels of difficulty and starting from the most difficult were vocabulary items that ranked respectively as first, with
13.95% success rate (consistent), second, with 16.28% success rate (integrated and outlets), fourth, with 23.26% success rate (queries), fifth, with 25.58% success rate (detect), sixth, with 27.91% success rate (contributions and clustered), and seventh, with 30.23% success rate (stationary).

Words such as “subsequent” (question 10) and “absorb” (question 14) ranked 10th giving the examinees a success rate of 37.21%. The word “dedication” (question 5) with a success rate of 39.53% ranked 11th. All the three words were classified as Difficult.

The word “narrows” (question 25), with a success rate of 51.16%, ranked 16th whereas the phrase “accounts for” (question 17), with a success rate of 65.12%, ranked 20th. Based on the students’ scores, both vocabulary items were Moderately Difficult.

Three vocabulary items turned out to be easy for the test-takers, namely, “statistics” (question 35), “hues” (question 41), and “founded” (question 7) with the first two terms ranked as 23rd at 79.07% and the third word ranked 24th at 81.4%.

Discussion

In the tabulation showing the overall ranking of the test items contrasting its difficulty against the rest of the items, a combination of six different types of questions forms part of what the respondents found to be Extremely Difficult. On top of the list is Vocabulary having only a total of six students responding to it correctly, i.e., the item was easy, the equivalent percentage of which is 13.95%. Based on Schraw and Roedel’s levels of difficulty, said question came out to be an Extremely Difficult one for 86.04% or 37 students of the entire TOEFL population for Quarter 2-2009. Still based on same band, the other types of questions the subjects found extremely difficult were Facts & Details, Tone & Purpose, Inference, Main Idea, and Organization although not necessarily in this exact order. This does not mean though that these types of questions were no longer classified again under any of the three other levels of difficulty, namely, Difficult, Moderately Difficult, and Easy. On the contrary, because of the huge number of questions for the entire Reading Comprehension Section, most of the question types were reclassified across the 50-item Reading Comprehension Section. To illustrate, although Vocabulary topped the rank within the Extremely Difficult level, the respondents found other Vocabulary test items to be Difficult, Moderately Difficult, or Easy depending on the degree of specificity of the question.

What appeared critical, in as far as vocabulary items were concerned, was that the examinees’ performance, generally speaking, posed as a reminder for students to pay more attention to their vocabulary, if only to better their scores. Evidently, with eight vocabulary items dominating the Extremely Difficult category, it may be safe to assume that poor vocabulary must have contributed to their not having answered correctly most other items albeit these items followed different formats and tested other reading skills. Steven Stahl, in his 1999 study,
categorically stated how strongly correlated reading comprehension and vocabulary are. Cynthia and Drew Johnson (2004) further asserted that “Limited vocabularies prevent students from comprehending a text.” This, however, is not surprising as it can be taken as a direct consequence of the students’ not being fond of reading texts in English.

With the findings considered unique to this specific group of students, it is nevertheless equally interesting to know what the general levels of difficulty each type of question had. Although all of the 50 items performed differently as the questions were mutually exclusive from each other, establishing and studying the general pattern the questions took is worth a look. As some question types appeared several times within the Extremely Difficult category, their sequencing was done based not on the frequency of the items, but on when each type appeared the first time within said category. As such, Vocabulary items topped the list as the most difficult item followed by Facts and Details. The third on the list was Tone and Purpose, fourth of which was Inference, fifth was Main Idea, sixth was Organization, and last was Referents.

While this study upholds the unquestionable assistance every language teacher can extend to language learners, the findings of this study nevertheless espouse the pedagogical merit of looking into the very specific circumstances surrounding language learning issues that are not always the center of attention in the everyday classroom.

In view of the current practices in most reading and TOEFL preparation classes, this study underscores the following points:

First, the subjects’ main difficulty was on vocabulary primarily because (1) their lexis was insufficient, and (2) they either had no or lacked knowledge of the fact that word meanings change in different contexts.

Second, questions in expanded format, especially those that required the respondents to use higher-order thinking skills such as synthesis and analysis, proved to be difficult as combining new ideas to form a new whole involved cognitive and meta-cognitive tasks they could not do with ease.

Third, inferential questions, generally speaking, were difficult to tackle not only because of the examinees’ poor vocabulary, but also because of their limited schema making them unable to make sense of what the texts indirectly said.

Fourth, making associations between a text and a related subject matter outside the passage was a difficulty which may be attributed to their poor world/background knowledge.

Just as the study offers its findings to be beneficial to some identified sectors in the field of language teaching, it also provides a number of suggestions that may be undertaken in exploring other language teaching-learning possibilities.

This study recommends the following undertakings relating to vocabulary enrichment: students need to expand their vocabulary both through explicit vocabulary instruction and sustained outside reading; students have to be taught about the “pervasiveness of contextual
variation in meaning” (Nagy, n.d.) and be trained how to recognize these changes so as to raise their level of awareness.

To address the students’ need for exercises that address the higher-level thinking skills, this study proposes that more synthesis and analysis exercises be given to the students to develop in them the ability to put ideas together to create a sound outcome. Teachers should demonstrate how to tackle a question in expanded format.

Given the fact that schema development or enriching the students’ background knowledge remains to be a challenge, this study supports the move to get students more exposed to texts with implied meanings. Teachers should help learners identify implied meanings in texts first, by demonstrating, and, second, by allowing them to try it on their own until they arrive at a desired level of reading writers’ implications. Furthermore, students should be more exposed to readings rich in world meanings aimed at deepening their world/background knowledge. Teachers should be able to demonstrate to them how world/background knowledge can be used to better understand various texts.

As there was insufficient data gathered, limitations were met which did not allow the researcher to study the learners’ test-taking strategies. This paper instead suggests that another investigation looking into said area through the use of verbal protocol be done in the future to further confirm results of this study: as propelled by the constraints encountered while gathering data, test items used were not authentic TOEFL questions, and, as such, some possible differences might have led to disparity in findings as opposed to using genuine TOEFL questions. A replicate study may be done in the future instead using, if possible, real TOEFL test items. With respect to the above recommendation, should a replicate study be done in the future, it is important to note that exactly the same testing conditions be observed so as to achieve an experience closest to that of, if not the same as, the actual TOEFL exam.

Lastly, this study recommends that findings and conclusions of this study be taken as something unique to the experience of the research participants and not an absolute representation of the entire PC population or of the MUIC students taking TOEFL as part of the university’s admission requirements.

In sum, this study confirms the importance of paying close attention to some very specific language difficulties second language learners encounter especially when said issues are left unattended outside the classroom context, providing learners with no support system they so need.

References


© 2011 Time Taylor International  ISSN 2094-1420
About the Author

Analiza Perez-Amurao obtained her AB-BSE in ELT from the Philippine Normal University in 1992. In 2006, she finished her MA in English Language and Literature Teaching at the Ateneo de Manila University. That same year, she obtained her Postgraduate Diploma in TESOL from RELC-Singapore. She currently teaches and acts as Coordinator at the Preparation Center for Languages and Mathematics at the Mahidol University International College, a leading state university in Thailand. An article she wrote on the theme “Innovation in Education” placed 5th in the recently concluded 2010 SEAMEO-Australia Press Award.
The Use of Study Strategies on Mathematical Problem Solving
Carlo Magno
De La Salle University, Manila, Philippines

Abstract
The study focused on differentiating the study strategies (deep, surface, and disorganization) on different cognitive skills in a mathematical problem solving test. Participants included 300 high school students from different high schools in the National Capital Region (NCR) in the Philippines. The participants were given the Students’ Study Strategy Questionnaire and a mathematical problem solving test. The One-Way Analysis of Variance (ANOVA) revealed that there were no significant differences on the usage of the three study strategies on the students’ mathematical problem solving across different cognitive skills. All study strategies were equally useful in solving different cognitive skills in mathematics. Surface approach was found to be functional for problems that require evaluation skills.

Keywords: Study strategies, mathematical problem solving, cognitive skills

Introduction
In the process of acquiring knowledge, individuals develop and make use simple to complex learning strategies to process information which is an integral part of mental development. In order for individuals to learn effectively, they use various methods and strategies in order to process information at a certain time frame (Sternberg, 2003). A study strategy or learning strategy is utilized depending on the preference and the way an individual approach a given problem (Kay & LeSage, 2009). According to Elliot, McGregor, and Gable (1999), there were three types of study strategies: (1) Deep processing, (2) surface processing, and (3) disorganization. Individuals who use deep processing analyze the deeper meaning of what is being studied (Craik & Lockhart, 1972). A person is considered a surface processor when one simply memorize and rehearse things that he/she is trying to learn (Veenman, Prins, & Verheij, 2003). Lastly, disorganization is seen when an individual is unable to organize one’s thoughts and keep a well-structured approach during one’s learning (Bandalos, Finney, & Geske, 2003). The three study strategies, or learning strategies are utilized in various ways depending on who used them and the situation. Most often than not, the use of these strategies were more clearly seen in the way students fulfilled their academic requirements. An important difference between a successful student and the student who is struggling is the way they go about learning (Protheroe, 2002). Students who are able to use study strategies effectively are able to be more successful academically. Study strategies have also served as tools to determine the strengths and needs of a certain student.

In a survey of literature done by Bandalos, Finney, and Geske (2003), it was determined that there were very few studies investigating the
connection of cognitive skills including problem solving and the use of study strategies to the success rates of students. An individual’s cognitive style influenced their performance and achievement in learning (Riding & Rayner, 1998). The ease by which an individual solved problems also depended on the study strategy employed. According to Leahey and Harris (1997), a problem occurs when there is a gap that separates a person from his goal. Problem solving is present in several aspects of life, from games to real life puzzles. Zaslavsky (1998) explained that if students were able to connect what they have learned inside the class to the events outside, they are able to better maintain and appreciate information.

Problem solving is also clearly seen in mathematics. In this subject, students are usually asked to comprehend a question, extract the necessary details and form a solution to get the correct answer. To obtain the proper answer, it is important for students to be able to make use of his metacognitive techniques.

The present study investigated the differences of deep, surface and disorganized strategies in solving mathematical problems. Most commonly, studies that used mathematics as dependent measures of study strategies are tested on a single array of domain and generalizations are made for mathematics in general. The present study further looked at the separate effects of study strategies on different cognitive skills in mathematics problem solving using Bloom’s taxonomy: Remembering, computation (understanding), applying, analyzing, evaluating, and creating (Byrd, 2002).

**Study Strategies**

A strategy is a method that would be able to provide some solution to a problem and give information (Best, 1999). Learners make use of strategies in order for them to be able to learn effectively. Study, or learning, strategies were defined by Weinstein and Mayer (1986) as behaviors or thoughts that a learner goes through during the learning process, and that can affect one’s encoding, storage, organization and retrieval of knowledge. According to Elliot, McGregor, and Gable (1999), there are three study strategies that students utilize. These three strategies were the deep processing, surface processing, and disorganization. Deep processing is referred to as the understanding of meanings and grasping connections between concepts. Surface processing, on the other hand, involves rote learning or simple memorization. Lastly, disorganization refers to the study strategy that did not follow any structure in learning.

Study strategies were important in understanding problem solving because individuals utilize several strategies in order to come up with a solution to a problem. It is important for an individual to be aware of what strategies to use to solve a given problem due to the fact that some strategies were able to generate answers easier if used correctly. When a person makes
a mistake in solving problems, they need to reorganize the information available to oneself and then proceed to attempt in solving problems (Catania, 1998). Like mentioned before, study strategies are used in gauging whether or not individuals will have an easy or hard time in solving various types of problems. However, it was said that not all strategies are useful, or considered the best (Alvermann & Moore, 1991; Anderson & Armbruster, 1984; Devine, 1991). On the other hand, Salovaara (2005) said that the proficient use of learning or study strategies included the easy shifting among the different study strategies and the application of its flexibly in compliance with the demands of the given task. Furthermore, Purdie, Hattie, and Douglas (1996) stated that learning a particular task required a different type of study strategy. One strategy could be used to solve problems better in some particular tasks while it could not be effective in other tasks.

The study strategies model by Elliot, Gable, and McGregor (1999) focused on three types of study strategies, namely deep processing, surface processing and disorganization. This model was used in the present study because of the appropriateness of the content of high school students.

Deep Processing. Deep processing is used when students understand the information that are being studied, while keeping in mind that the small details they received are aids in solving the problem successfully (Biggs, Lai, Tang, & Lavelle, 1999). A deep processor is somebody who look profoundly into the subject matter, analyze them thoroughly and then tries to apply previous knowledge to accomplish his task (Entwistle, 2002). Deep processing is important in problem solving because it allows an individual to use previously learned concepts to find solutions for new problems, especially if the previously solved problems and the new ones have similar characteristics.

For some students, deep processing is strengthened by the fact that it entailed more time in pursuing their own interest since they go through good time management (Prosser & Trigwell, 1999). It can also be a positive experience in education that leads confidence in one’s ability to understand and succeed. In addition, deep processing can be seen when one was determined to do well and be mentally occupied in their academic work for the reason that students would be intrinsically curious in the subject matter.

Teachers also played a part in influencing students on whether the latter will use the deep approach as their study strategy or not. When teachers are interested in the subject matter, it would have an impact in the student’s choice of study strategy. Teachers might also use tests that required thoughts and ideas be used together (Prosser & Trigwell, 1999). Encouraging students to be active in class and correcting students’ wrong beliefs would greatly help students to properly understand the subject matter. Likewise, if the teachers were fair in assessing learning and allowed students to commit mistakes without any possible penalty, such as being embarrassed in class, as well as rewards, such as favoritism, students were
observed to be more eager to learn and therefore proceeded to use the deep approach.

In the study of Marton and Saljo (1976) regarding deep and surface approaches to learning, the results showed that deep processing was actually the study strategy that most educators would like their students to develop. Deep processing helped students to develop more elaborate and stronger memories. It was said to lead to good understanding and as well as the achievement of better grades. However, any study strategy, when harmonized with the learner's cognitive strengths, was said to produce positive results and attitudes (Alesandrini, Langstaff, & Wittrock, 1984).

In a study by Lockhart, Lamon, and Gick (1988), they argued that participants should engage in higher conceptual processing for the prior presentation of information to transfer when solving the problems. They argued that the problem should be seen as a puzzle and not as knowledge retrieval, for which the skill of capacity would be needed. This meant that deep processing was needed to develop practical thinking in the sense that similar problems could be solved through the steps and reasoning gained through deep processing.

In another study made by Entwistle (2002), it was shown that the deep approach was more common in departments whose students’ rates were in the good teaching and allowed freedom in learning. For the students, good teaching was assessed in the areas of level, pace, structure, explanations, enthusiasm, and empathy. On the other hand, freedom in learning was shown by having the teachers extend their style of teaching in a more creative way. In this way, students were more encouraged to learn and have self-reliance. On the contrary, giving students a heavy workload or tests emphasizing an accurate copy of the lecture were more likely to lead into the use of surface approach.

In relation to the achievement goal theory, deep processing had a positive relationship with the mastery goal orientation (Ford, Smith, Weissbein, Gully, & Salas, 1998). In this type of orientation, one focused on goals such as learning as many as possible, overcoming challenges, and increasing one’s level of competence (Wolters, 2004). Individuals who followed the mastery goal orientation utilize the deep approach, one requiring a lot of cognitive effort but lead to better understanding (Ford, et al., 1998).

Bernardo (2002) was able to mention in his study some alternatives to the traditional method of teaching school mathematics. These alternative methods were said to be focused on the active participation of the student, leading to better understanding and not simply memorization. These alternative methods and the deep processing approach were similar due to their focus on letting a student learn the concepts of a subject matter. It was, therefore, believed that students have fully learned mathematics when they were able to successfully explain and prove their answers.
It has been accepted that deep processing was the excellent choice in study strategies for one to obtain a high performance in mathematical problem solving. In view of the fact that mathematical problem solving required people’s organized framework to understand different topics, deep processing was to be used, given that deep processing was known to be the approach wherein one will look for meanings and understand the concept being studied and be able to relate it with other experiences (Penger, Tekavcic, & Dimovski, 2008). Being able to connect deep processing and mathematical problem solving, one can solve mathematical problems easily and accurately, and would have higher performance (Busato, Prins, Elshout, & Hamaker, 1998).

**Surface Processing.** In the surface processing strategy, students study information mainly by what they could observe (Craik & Lockhart, 1972). Bernardo and Okagaki (1994), Baroody and Ginsburg (1986), and Reed, Dempster and Ettinger (1985) observed that several students who were successful in solving a particular type of mathematical problem had difficulty in solving problems that might only have a slight difference from the ones that they were used to solving every time. Students who utilized the surface type of learning focused their attention on data that they felt were important and proceeded to memorize that specific information (Biggs, 1990). Surface processing can also lead to shallow retention of material in the examinations.

According to the Matthews (2001), surface processing does not enhance understanding in the long-term retention of information and knowledge. He compared characteristics and factors that brought about the use of deep and surface approach. Surface processing caused high anxiety in students. They would rather resort to the use of surface processing either because of too much workload or because they were not given enough time to finish their work. In addition to this, students may have a doubtful view of education, believing that factual recall is the process that best and was required for tests. Students were also evaluators of their own learning, and the evaluations are said to often follow the patterns given by their professors so as to patterns in solving mathematical problems. Surface processors may not involve with learning how to do their tasks the right way and may have not really understood why they did certain processes, which led to more usage of the surface strategy in their studies (Matthews, 2001).

Having interest in the subject matter for both students and the teacher was important in the development of an approach. When teachers looked uninterested and allowed the students to be passive, surface processing had a higher chance of being developed. Aside from this, when teachers imposed that students be always active in class or made low expectations and discouraging statements, students are likely to be less motivated to study, thus the student would study just for the sake of passing. Surface processing results to teachers who keep on rushing lessons so that they would cover a lot
and also overemphasized things, such as the syllabus, will inhibit high quality learning (Entwistle, 1998).

The surface approach to learning had a positive relationship with performance goal orientation. Individuals who have a performance goal perform a task better than others, making them focus on performing quickly and efficiently. In this case, students resorted to a more shallow approach when it comes to learning (Ford et al., 1998).

Traditional teaching methods in mathematics usually involved repetition of the processes being indicated during the class (Brown, Cooney, & Jones, 1990). Bernardo (2002) mentioned that in this setting, the teacher and the textbooks becomes the main source of information and them alone. In this setting, students were said to have learned their lessons when they were able to obtain the right answers by simply following the instructions given to them. When students only followed their teachers without thinking on their own, students were then adapting a surface approach to learning. They are focused on simple rote learning and repetitions as parts of their study strategies to learning. There is endorsement of surface processing because in reality, tests given to students did not require a deeper level of learning in order for the students to pass their examinations (Sankaran & Tung Bui, 2001).

In another perspective, the study conducted by Rhem (1995) showed that the more the students advanced in their academics, the tendency for them to use surface processing also increased. He mentioned in his paper that the use of surface processing might be due to the traditional teaching styles being used in schools. He also listed a couple of other factors that might have been responsible for students to resort to surface processing: (1) an excessive amount of material in the curriculum, (2) relatively high class contact hours, (3) an excessive amount of course material, (4) a lack of opportunity to pursue subjects in depth, (4) a lack of choice over the method of study and over the subjects, and (5) a threatening and anxiety provoking assessment system.

The emotional consequence to surface strategy was explained in the study by Entwistle and Waterston (1988). They used interviews and questionnaires to determine the comparative strength of deep and surface approaches. The findings showed that deep approach was constantly linked with academic interest in the subject matter and that the individuals who used this kind of approach resulted to be self-confident while surface processing was constantly associated with anxiety and fear of failure.

The study conducted by Prosser and Millar (1989) showed that students who used the deep approach changed their perception of technical materials in ways which the lectures required. Students who relied on the surface approach did not develop the necessary understanding of the technical materials required by the course. Because of this lack of a solid
foundation, students who used surface processing had more problems as the course progressed.

Surface processing only used memorization to study, while mathematical problem solving required a certain level of analysis and understanding (Hong & Aqui, 2004). With this, one can see that surface processing do not help individuals maintain different mathematical operations in problem solving because they do not properly understand the concept behind the problem because the individual simply memorized the processes. When teachers arranged the assessments given to the students in a different manner, the latter might not be able to answer the questions appropriately, due to the fact that they only memorized the material and did not understand the concepts behind the material. Schoenfield (1988) mentioned that some students were believed to have understood various concepts when they were able to replicate the exact things that their professors did. However, Schoenfield (1988) stated that there were students in reality who simply went through their studies in a mechanical way, without truly analyzing the things that they studied.

The study by Magno (2009) using Filipino students provided a different perspective for deep and surface strategies in learning. He found in his study that there are equal opportunities for the use of deep and surface strategies when related to ability and metacognition measures. Other studies made by western researchers focus on the advantage of the deep strategy and less for the surface strategy. However, for Filipino students, they see surface strategies as functional in their learning. This result challenges existing models about the consequences of deep and surface strategy use.

**Disorganization.** In disorganization, an individual has a difficult time in building and/or maintaining a stable and organized way of study that is used for learning (Bandalos, Finney, & Geske, 2003). According to several researches (Bhaskar & Simon, 1977; Chi, Feltovich, & Glaser, 1981), effective problem solving depend strongly on both the nature and organization of knowledge that is within the individual's accessibility (Bransford, Sherwood, Vye, & Rieser, 1986). Given this case, individuals who use the disorganization approach have a hard time concentrating and analyzing problems.

There are only a small number of studies that focused mainly on disorganization as a study strategy. Despite this, the small amount of research showed that disorganization and performance had a negative relationship (Entwistle & Ramsden, 1983). Another point was that disorganization and test anxiety had a positive relationship (Bandalos, Finney, & Geske, 2003). Elliot, McGregor, and Gable (1999) stated that having an unorganized study strategy cause a student to become unprepared for an examination. Unpreparedness then lead to anxiety when the particular student was to be evaluated by other people (Elliot, McGregor, &
Gable, 1999). In the same way, Al-Emadi (2001) showed a negative correlation in mastery goal and performance goal to disorganization, and a positive relationship to performance avoidance goal.

When individuals use disorganization as a strategy, they have difficulty in maintaining a stable and structured way of learning (Bandalos, Finney, & Geske, 2003). Given that mathematical problem solving use highly structured and organized thinking, an individual who use this kind of approach had difficulty analyzing problems presented to them. In the study of Elliot, Gable, and McGregor (1999), having an unorganized study strategy was a problem for students for the reason that they might be caught unprepared for assessments which involved mathematical problem solving.

**Problem Solving**

These are instances when individuals especially students are required to show their ability to solve problems (Groome, Dewart, Esgate, Gurney, Kemp, & Towell, 1999). Problem solving is a skill that is considered to be important. It is a skill that could be used throughout the course of one's life not simply in academics but even in getting or retaining a job position. It is considered as the outcome of teaching a content of an activity after a math skill lesson. Although most of the time, problem solving aimed at something as its goal, it does not necessarily mean that this end goal was always achieved (Leahey & Harris, 1997). An investigation done by Neef, Neles, Iwata, and Page (2003) showed that individuals’ who solve problems effectively posses the ability to transform the words in a problem into their necessary symbolic counterparts and this was significantly related to one’s ability in problem solving. According to the Miller, Hall, & Heward (1995) individuals should be reminded that they cannot arrive at perfect solutions all the time; however, one can create the best possible decision with the information given at hand. Ong, Liao, and Alimon (2009) even mentioned that there were times when difficulties in problem solving arose when individuals jumped immediately in finding solutions to the problems simply because they did not properly identify first the elements needed in the problem.

A number of early studies regarding the cognitive processes involving problem solving had already been done and this was narrated by Groome, et al. (1999). Oswald Kulpe was one of those who took an interest in researching about thought processes. Next, were the behaviorists who claimed that thought processes were due to observable behavior. One of these behaviorists was Thorndike in 1898 who said that it was through the process of trial and error that problem solving was done. Gestalt psychologists agreed that trial and error may play a part in problem solving; however, they still continued to conduct several well-known studies regarding problem solving. Köhler, one of the founders of Gestalt psychology, was able to study
problem solving skills while observing apes. Bananas were hung from the ceiling and the ape was observed as it tried to reach for the fruits. After some futile attempts, Köhler discerned that the ape seemed to have insights, because it decided to stack crates in order to make a stairway and reach the fruits. For the Gestalt psychologists, insight, or the sudden understanding of a problem because of reorganization and restructuring of the elements of the problem, was an important constituent for successful problem solving. Insight, for Gestaltists, was an extraordinary method of thinking that was unlike the usual linear processing of information (see Sternberg, 2003).

During the 1960's, a study about human problem solving was headed by Herbert Simon (Dunbar, 1998). In Simon et al.'s investigation, he made use of complex problems which involved several characteristics that might lead to a solution. Next, concurrent verbalizations from subjects were utilized to recognize mental operations, strategies, and representations that individuals used for problem solving. Lastly, Simon et al. made use of computer programs that mimicked human processes in problem solving.

Solutions to problems do not come to a person easily. On the other hand, there were two usual ways being utilized for solving problems. These methods were referred to as algorithms and heuristics.

**Algorithms.** Algorithms are solutions or strategies being used that were sure to produce solutions (Leahey & Harris, 1997). According to Leahey and Harris (1997), algorithms were best for well-defined highly-structured problems. Algorithms were also commonly used in areas pertaining to the use of computers and other types of technology.

It has been said that algorithms may not always be utilized because of well-structured and ill-structured problems (Best, 1999). Ill-structured problems were defined as those problems without clear solution paths (Sternberg, 2003). Because of this, Best (1999) mentioned that there was no possible way for people to establish an algorithm to solve the problem. Well-structured problems, or those problems with clear solution paths (Sternberg, 2003), may make use of algorithms, but, according to Best (1999), one might be disappointed due to the fact that well-structured problems were usually large. In other words, algorithms may still not be able to fully embrace a lengthy solution to a problem despite it being able to help people in solving simple well-defined problems.

Examples of algorithms used in well-defined or well-structured problems would be geography and mathematics (Sternberg, 2003). The mathematical concept of multiplication could be considered as a process that made use of algorithms, because if one knew the procedure of multiplying numbers, one will surely be able to find an answer to any multiplication problem as long as one followed the given set of rules (Reed, 1996).
**Heuristics.** It was not possible to always have a ready answer for every problem. Therefore, the use of algorithms was usually exchanged for the use of heuristics. Heuristics were problem solving strategies that involved the use of experience and gut feelings (Leahey, & Harris, 1997). This type of problem solving strategy was usually used, albeit success was not warranted (Reed, 1996). Heuristics was utilized mainly in many real life situations because there may be none to little algorithms available for one’s use (Groome, et al., 1999).

Examples of heuristics would be forming subgoals and using analogy. The first of the two, forming subgoals, was when a problem was divided into several parts (Reed, 1996). A subgoal was defined by Reed (1996) as a goal that served as a solution for a part of the initial problem. This type of heuristic allowed an individual to avoid erroneous paths to the solution due to having an idea of how the solving should be done (Reed, 1996). The drawbacks for this type of heuristic would be that not all subgoals were obvious in problems, and that there were times when subgoals triggered uncertainty instead of giving the solver an easier time with one’s work (Reed, 1996).

Another example of heuristics would be the use of analogy. The usage of analogies called for an individual to make use of other analogous problems in order to solve a particular one (Reed, 1996). In order for an individual to be able to do this, one should first recognize the similarities between or among the previous problems and the problem presented at hand. One should also be able to recall the proper solutions to the previous problems and be able to apply it to the present problem (Reed, 1996).

Solving problems was not limited to the usage of algorithms and heuristics; other researches had also focused on the role of creativity in problem solving as well (Sternberg, 2003).

**Mathematical Problem Solving**

Mathematics is one of the subjects that require cognitive processes and is highly intellectual in nature (Hong & Aquí, 2004). It involved solving simple equations to complicated ones. Despite this, mathematics is claimed not only solving of problems with the use of complicated formulas, but to be a stepping stone on how one should think and apply what one has learned in real life (Naglieri & Gotling, 1995). Mathematics was also an excellent field to determine the success and failure rates of the students depending on the study strategies that they utilized.

Problem solving was not limited to the events occurring in the real world. There are times when problem solving was linked to mathematics because it is an essential part of the latter (March & Cooke, 1996). Mathematical problem solving was said to be a transfer challenge requiring individuals to develop schemas for recognizing novel problems as belonging to
familiar problem types for which they knew solutions (Fuchs, Fuchs, Finelli, Courey, & Hamlett, 2004). Individuals may also have to learn to make use of synthesis, which was defined by Riley and Greeno (1998) as putting together various elements of a problem, to come up with other solutions.

Mathematical problem solving was said to be pioneered by George Polya (Higgins, 1997). He was able to develop a four-phase model of the problem-solving process. This model involved the following steps: (1) understanding the problem, (2) devising a plan, (3) carrying out the plan, and (4) looking back. These steps could be explained through Young’s (1924) similar problem-solving model. Step 1, understanding the problem, involved an individual getting a clear idea of what information are being asked in the problem. The next step involved the planning stage where an individual decides which information will be useful in his search for an answer and what strategy one must use to get the desired result. In the third step, the person will try to implement his plan. If his first plan does not succeed, he continues to implement other plans until he is able to succeed. The last step, looking back, involved taking a step back and checking whether or not the result satisfies the data being asked in step one.

Some studies done regarding students’ performance when it comes to mathematical problem solving were done by Cote and Levine (2000), and Botge (2001). In Cote and Levine’s (2000) study, a student’s attitude towards mathematics was a factor in their performance. A positive attitude towards math allowed an individual to be more successful at solving mathematical problems, compared to those who had negative views about the subject matter. In Botge’s (2001) study, he found that when individuals are able to understand the basic mathematical concepts, achievement was also high. He also found that an individuals’ skill in computing was weakly correlated to that of one’s achievement. Benedetto-Nasho and Tannock (1999) found that an individual’s errors in computations was less likely the reason with regards to poor mathematical problem solving performance, when compared with the amount of errors in the problem representations.

Strategies used in problem solving had been one of the main focus of research regarding mathematics education (Schurter, 2002). Mathematical operations and problem solving make use of algorithms, a type of strategy used by people and which were sure to generate solutions for given problems (Leahey & Harris, 1997). Sternberg (2003) mentioned that a type of mathematical concept that could be considered as an algorithm would be multiplication. Mathematics was a subject matter that involved a number of formulas and equations, and if properly used will yield the results needed without fail. Unfortunately, the field of mathematics requires a lot of thinking (Hong & Aqui, 2004) and was considered as one of the most difficult subject matters. Koller and LeSage (2009) mentioned that the problem with a lot of students was that when it came to mathematics, these students believed that certain problems were unsolvable if they were not able to detect
a solution for the problem at once. In mathematical problem solving, one needs several cognitive skills such as identifying the elements, computing, analyzing the problem, synthesizing and evaluating. In Bloom’s Taxonomy, problem solving is said to consist of six major categories that start from the simplest to the most complex skill. These categories could be considered as degrees of difficulty and individuals must master the first skill before going to the next. The six cognitive skills are remembering, understanding, applying, analyzing, evaluating, and Creating (Byrd, 2002; Pohl, 2000).

The Present Study

In the present study, the three study strategies of deep, surface, and disorganized are differentiated on solving mathematical problems. There is a string of evidence in previous studies that deep processing was the best study strategy to use in order to perform highly in mathematical problem solving. This was explained by Hong and Aqui (2004) when they said that deep processors got to grasp the entire concept behind various mathematical equations. Moreover it was explained by Busato, Prins, Elshout, and Hamaker (1998) that by having integrated different concepts into one's memory, a deep processor could easily utilize previously learned concepts into new and unfamiliar situations. Since mathematical problem solving was a transfer challenge that required people’s schema (set of beliefs that provides an organizing framework that would help us to understand an event, topic, or person) to recognize creatively made problems (Fuchs, Fuchs, Finelli, Courey, & Hamlett, 2004), deep processing can be a strategy that might help problem solvers to have a highly desirable output in solving a problem. On the other hand, surface processing give individuals optimal performance at a point in time. However, surface processing does not help individuals retain the method in solving mathematical problems which required a certain level of analysis and understanding (Hong & Aqui, 2004). This implies that without properly understanding the methods involved in mathematical problem solving, an individual will not be able to process the problems once it was different from the ones he memorized. The use of disorganization as a strategy was said to lead to the difficulties of a learner in developing a structured and organized approach in solving problems (Bandalos, Finney, & Geske, 2003). With mathematics being a highly intellectual subject matter, a student with no firm foundation for his/her way of learning will have difficulty in putting information he/she had gathered into actual computation or practice. In relation to test anxiety, having a disorganized study strategy could catch a student unprepared during examination.

The study focused on the different study strategies used on different cognitive skills on mathematical problem solving. It particularly tried to determine which study strategy would best facilitate effective learning in students. High school students are classified according to their dominant
study strategy and then a mathematical problem solving test was administered.

Method

Participants

The participants were 300 high school students from two different schools in Manila, Philippines. These two high schools were chosen for their string program on developing high achievers in Mathematics. The participants were fourth year high school students, both male and female students belonging to the 16-18 age group. High school students were appropriate because of the uniformity of the mathematics curriculum and competencies developed for high school in the Philippine Educational system.

Cluster sampling was then used to enable the proponents to get the 100 representative samples from each of the three learning strategy piles. In the present study, cluster sampling was done through the random selection of sections. The students were categorized according to the study strategy (deep processing, surface processing, disorganization) when solving math problems. Their scores were determined through the Student’s Study Strategies Questionnaire (Elliot, Gable, & McGregor, 1999).

Instruments

Student’s Study Strategies Questionnaire (SSSQ). The SSSQ was used to measure the study strategy used by the participants (Elliot, McGregor, & Gable, 1999). The questionnaire consisted of 15 items which measures three subscales: Deep processing, surface processing, and disorganization. There were five items that measured each of the subscales. The scaling technique used was a 7-point linear numerical scale. Principal-Components Factor Analysis with varimax rotation was used to uncover the three-factor solution of the item. The three factors, deep processing (eigenvalue=2.62), surface processing (eigenvalue=2.10) and disorganization (eigenvalue=3.31), accounted for 55% of the total variance. Past studies (Elliot & Church, 1997; Elliot & McGregor, 1999) have presented evidence for the construct validity, predictive validity and reliability of these measures.

Mathematical Problem Solving Test. The Mathematical Problem Solving test was constructed to measure the problem solving ability of the students. There were 25 items included in the test that covers high school lessons in mathematics. The coverage of the test included fractions, factoring, simple algebraic equations, and various word problems. These content were based on the learning competencies of the Department of
Education in the Philippines. The items were only limited in measuring identifying, computing, analyzing, creating and evaluating cognitive skills based on the revised Bloom’s taxonomy (see Pohl, 2000). The items were reviewed by a professor in mathematics in a University and a high school teacher teaching mathematics. The mathematical problem solving test was revised based on the review. It was pre-tested among high school students in a private high school to determine the length of time needed by students in answering the entire test. The test had an internal consistency of .84 based on the KR #20. The overall mean of the test performance of the participants in the pilot test was 16.23 with a standard deviation of standard deviation of 5.45. This showed that scores of 17 to 25 were high in problem solving and a score of 15 and below were below average. A standard deviation of 5.45 meant that the individual scores are dispersed. The difficulty of an item was based on the percentage of people who answered it correctly. The index discrimination revealed that there were no marginal items as well as bad items; however, 84% of the items were very good, 2% were good items and 2% were reasonably good items. It was also found out that there were no difficult items presented, although 72% of the items were average and 28% were easy.

Procedure

Approval letters were given to the different high schools to gain permission for the administration of the measures. The purpose of the study was explained and how the school can benefit from the study. The measures were administered during the participants’ mathematics classes.

The participants were informed that they will be answering a mathematical problem solving test and the study strategy scale. Instructions for answering the different tests were given orally to the students before they were asked to answer the Students’ Study Strategy Questionnaire (SSSQ). In the SSSQ, the participants were asked to choose the most appropriate response to each item on a scale of 1 (not at all true of me) to 7 (very true of me). They were told that they should not spend a long time on an item because there was no right or wrong answer. The participants were assured that their answers will be kept confidential. After finishing the SSSQ, the students were asked to take the mathematical problem solving test. They were instructed to read the sample items first and the instructions on how to answer the test. The students were then further instructed to shade the letter corresponding to the answer of their choice. They were asked not to write anything on the questionnaire since answer sheets were provided. Further questions were not entertained during test taking, except for questions which involved clarification of test questions. The students were given 90 minutes to complete all the questionnaires. After all the students have completed the test, they were debriefed about the purpose of the study.
Data Analysis

The participants were classified to their dominant study strategy by comparing the scores of each students across the three subscale of the SSSQ. The highest score among the three subscales were considered as the dominant study strategy of a participant. Students are then classified according to the study strategy that they belong. The study strategy was treated as a categorical variable in the present study.

The mean and standard deviation was obtained to determine the performance level of the students in mathematical problem solving. The One-Way Analysis of Variance (ANOVA) was used to determine whether there are significant differences among the use of deep processing, surface processing and disorganization in their mathematical problem solving.

Results

The One-Way ANOVA was used to determine if there were any significant differences among the three learning strategies (deep processing, surface processing and disorganization) on problem solving across its five components (identification, computation, analysis, synthesis and evaluation).

Table 1
One-Way ANOVA Summary Table

<table>
<thead>
<tr>
<th></th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study strategy</td>
<td>27.88</td>
<td>2</td>
<td>13.94</td>
<td>0.773</td>
<td>0.462503</td>
</tr>
<tr>
<td>Error</td>
<td>5156.38</td>
<td>286</td>
<td>18.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

There was no significance obtained in testing for the differences among students who used deep processing, surface processing, and disorganized study strategies on their problem solving as a whole, p=0.46. This means that whatever study strategy a student used, it did not make any difference in their problem solving scores.
The mean scores of the students who used disorganization ($M=16.12$, $SD=4.15$), deep ($M=17.24$, $SD=4.38$) and surface strategies ($M=17.21$, $SD=4.13$) on problem solving were almost the same. There were no significant differences among the three learning strategies when it came to problem solving in general.

The three study strategies for each cognitive skill of the mathematical problem solving test was also compared. Problem solving was categorized into the five cognitive skills, namely identifying, computing, analyzing, creating, and evaluating.

### Table 2

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>$SS$</th>
<th>$df$</th>
<th>$MS$</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying</td>
<td>1.05</td>
<td>2</td>
<td>0.525</td>
<td>0.493</td>
<td>0.61</td>
</tr>
<tr>
<td>Understanding</td>
<td>5.28</td>
<td>2</td>
<td>2.64</td>
<td>1.75</td>
<td>0.17</td>
</tr>
<tr>
<td>Applying</td>
<td>2.71</td>
<td>2</td>
<td>1.36</td>
<td>0.86</td>
<td>0.42</td>
</tr>
<tr>
<td>Analyzing</td>
<td>1.41</td>
<td>2</td>
<td>0.71</td>
<td>0.36</td>
<td>0.70</td>
</tr>
<tr>
<td>Evaluating</td>
<td>11.05</td>
<td>2</td>
<td>5.52</td>
<td>3.60*</td>
<td>0.02</td>
</tr>
</tbody>
</table>

* $p<0.05$

In testing for the significant difference among students who used deep processing, surface processing, and disorganized study strategies, no significance was found for identifying, understanding, applying, and analyzing. There were no true differences obtained with the usage of any among the three study strategies on the students’ problem solving across the first four skills. A significant difference of the three skills was only found for

**Figure 1. Line Graph of Problem Solving Mean Scores**

The graph shows the mean scores of students using different learning strategies on problem solving. The mean scores are as follows:

- Disorganization: $M=16.12$, $SD=4.15$
- Deep: $M=17.24$, $SD=4.38$
- Surface: $M=17.21$, $SD=4.13$
evaluating. As shown in Figure 2, the mean scores of the students who used surface strategy ($M=3.97$, $SD=1.20$) was significantly higher than those who used disorganization ($M=3.27$, $SD=1.30$) and deep strategies ($M=2.76$, $SD=1.16$) in evaluation with a medium effect size ($\eta=.22$). This means that students who used surface strategy were better on the evaluating items of the problem solving test. The students who used surface processing need not think too deeply due to the fact that they can easily check and see if the patterns in the choices match the ones that they memorized before.

![Figure 2. Line Graph of Problem Solving Mean Scores](image)

**Discussion**

The present research tested the differences among the three learning strategies (deep processing, surface processing and disorganization) proposed by Elliot, Gable, and McGregor (1999) on different cognitive problem solving. It also investigated the relationship between metacognition and problem solving. This chapter presents a discussion about the findings of the study.

**Deep Strategy**

The students who made use of deep processing scored an average of 17.24 in the Mathematical Problem Solving Test. It was found that the use of deep processing is not significantly different with surface and disorganized on the students’ problem solving skills in general. This was determined by the fact that the scores for deep processing and surface processing had a variance of only .03 points. A study by Alesandrini, Langstaff and Wittrock (1984) showed that learning was often correlated with capabilities not matching the study strategy used. As such, the results demonstrate the fact that the study strategies used by the students had minimal impact to the final scores in a test administered to determine the extent of their learning.
Surface Processing

The students who made use of surface processing obtained an average score of 17.21 when they took the mathematical problem solving test. It was found that the use of surface processing was not significantly different with the deep and disorganized on the students’ problem solving skills in general. Students who made use of surface processing did not necessarily perform lower than those who used deep processing or disorganization as compared with previous studies. The individuals who used surface processing had almost equal mean scores of those who used deep processing. This supports the results of Purdie, Hattie and Douglas (1996) that rote learning could be useful to some learning tasks and could lead to better understanding and generate acceptable solutions (Salovaara, 2005). However, the information that could be obtained from rote learning may not necessarily be sufficient to raise overall scores in problem solving. Instead, surface processing enabled students to have scores comparable to that of deep processing. Both deep and surface strategies contribute to the overall problem solving performance but they did not constitute the whole of the learning process critical to mathematical problem solving.

Disorganization

The students who made use of disorganization had an average score of 16.12 on the Mathematical Problem Solving Test. It was found that disorganization, when utilized in problem solving produced the same levels of benefits to the overall problem solving performance. Furthermore, the students have used disorganization as a study strategy, but these individuals did not rely on a single strategy during learning. This is especially true since Bhaskar and Simon, (1977); and Chi, Feltovich, and Glaser (1981) found that effective problem solving is a result of the nature and the organization that was accessible to the student at that given time. As such, while these students relied primarily on disorganization as their study strategy, it was feasible that they have also employed surface processing at one time or another because this strategy is not as directed as compared to deep and surface.

Differences among Study Strategies on Problem Solving

The usage of the three study strategies were found to have no differences on their roles in enhancing a student’s learning skills, especially in mathematical problem solving. This supports the findings of Alvermann and Moore (1991), Anderson and Armbruster (1984), and Devine (1991) that there was no single type of study strategy that would best empower a student
when it came to learning different tasks. They consistently explained that the learning strategies that students used depend not only with the type of student who uses it, but also in the situations where learning was going to take place (Alvermann & Moore, 1991; Anderson & Armbruster, 1984). It also supports the explanation of Devine (1991) and Salovaara, (2005) that skillful use of study strategies involved shifting smoothly among different study strategies and applying these strategies flexibly in accordance to the demands of the task given.

The results of the present study do not support the claims of Elliot, Gable and McGregor (1999). They indicated that there are differences in terms of an individual’s academic achievement which favors the use of deep strategy. The results showed that deep strategy is not more advantageous than the other strategies in aiding a student’s learning process when solving word problems. In the case of this study, the data was not able to show sufficient evidence to prove that those who used deep processing will indeed have better outcomes when it came to their problem solving abilities. This might be because the system of teaching in high school may not have fostered the students’ abilities with regards to independent thinking. This supports the explanation of Magno (2009) that Filipino students see the function of a surface approach equally adaptable and useful as the deep approach to learning. Contextual factors in performing a task require rote memorization of information in problem solving which students find useful.

**Differences among Study Strategies and the Five Cognitive Skills**

The present study expanded the existing body of reviews by looking at the effects of the different study strategies on the five different cognitive skills used in problem solving. These five cognitive skills were identifying, computing, analyzing, creating, and evaluating.

**Identifying.** Deep processing, surface processing, and disorganization were found to have no significant differences when used on the first component of mathematical problem solving which was identification. Identification in problem solving involves questions that make students to look for the values being asked. The difficulties of problem solvers was not in detecting the elements in the problem, but by jumping straight into trying to figure out the solution without identifying the variables concerned (Kay & LeSage, 2009.). The findings did not support the claims of Riding and Rayner (1998) that a person’s cognitive styles influence one’s academic achievement and performance.

**Computing.** The three study strategies were found to have no significant differences on the computation of problem solving. This means that whatever type of study strategy a student uses, it does not affect how
well a student controls his or her computational skills. This was explained by Benedetto-Nasho and Tannock (1999) when they said that poor mathematical problem solving abilities may be caused more by the mistakes in representing the problems and less by the errors committed in relation to actually solving the problems. This finding was supported by Naglieri and Gotling (1995) when he found that an individual’s skills in computing have a weak correlation with the individual’s achievements. He also found that it was essential for students’ to have an understanding of the basic mathematical concepts in order to have higher achievements. However, despite the results of Schoenfeld (1988) study, computation was still believed to be an essential component of mathematics. This means that although previous research was found that computation was not directly related with a student’s success in academics, it was still an aspect that was explored to see if it was affected by the way a student studies in preparation for an examination in mathematics.

**Analyzing.** Deep, surface, and disorganization were found to have no significant differences on the third component of mathematical problem solving which is analysis. Wilson (1993) explained that students are believed to understood mathematical concepts as long they simply replicate the exact routine on how their professors have gotten the correct answers to a problem. He also said that in reality, the students may only have mechanically gone through the steps to obtain the final answer. This means that a student’s analytical skills in mathematical problem solving may not always rely on the type of learning, or study, strategy that they use depending on the nature of the problem given to them. The act of solving a mathematical problem may involve little to no analysis at all when the same pattern of problem is given (Wilson, 1993). Because of the simple repetition of the Professor’s chosen method, students no longer relied on the deep method of disassembling a problem, instead skipping ahead to the formula. As such, no matter what method of study is used, the students displayed no variation in their performance.

**Creating/Synthesizing.** The results of the present study showed that deep processing, surface processing, and disorganization were found to have no significant differences on the synthesis of mathematical problems. Synthesis required only the conscious act of putting all the information or the elements in the mathematical problem together, then start forming a new solution or equation, among others (Tall, 1991). The use of any of the strategies allowed the student to pick elements needed to be put together in arriving with a solution. As long as the students were able to create their own solutions, there was no variation in their performance in problem solving. The study strategies employed do not affect this process. This finding was supported by Botge (2001) when she suggested that one way that was used to synthesize ideas was simply by making an outline or creating a list of the
elements, then putting all the elements and their relations together afterwards. The study strategies granted the students a mental outline of the solution, which the students now then used as a basis for the final answer that they came up with.

**Evaluating.** The strategies deep processing, surface processing and disorganization were found to have a significant difference on the fifth component of mathematical problem solving which was evaluation. It was found that students who used surface processing were able to evaluate their solution and/or answers well compared to those who used deep processing and disorganization. Since evaluation looks for the right judgment in order to get the right solution or answer, students tend to use their judgment on the best way to arrive with the correct answer. The use of the surface strategy in the in problem solving looks at the value of tacit knowledge which is part of rote memorization (Biggs, 1999). Rote facts and memories are used to scaffold students to solve problems that require evaluation. When evaluating, students find usual patterns on relating two factors: Know-how abilities and problem solving/thinking. Surface processors who made use of memorization also used the patterns to their advantage in evaluation what they have done is correct.

**References**


About the Author

Dr. Carlo Magno is presently an associate professor at the Counseling and Educational Psychology Department of De La Salle University, Manila. Most of his researches and publications focus on learning strategies, self-regulation, metacognition, language learning, and student performance. Further correspondence can be addressed to him at carlo.magno@dlsu.edu.ph
Confirmatory Factor Analysis of the Academic Procrastination Scale
Romel A. Morales
*University of Eastern Philippines, Northern Samar*

**Abstract**

This study verified the construct validity and reliability of the Academic Procrastination Scale (APS). Undergraduate students (*N* = 1153) from 5 different colleges at the University of Eastern Philippines completed the APS. Confirmatory factor analyses (*CFA*) indicated that a three-factor structure of the APS provided a good fit to the data. The APS dimensions and items had moderate to high internal consistency reliabilities indicating the usefulness of the scale in measuring the dilatory behavior of students especially in college setting. These results support the utility of the scale for research and in theory development of academic procrastination.

**Keyword:** Confirmatory Factor Analysis; Validity; Reliability; Structured Procrastination, Unstructured Procrastination, Non-procrastination

**Introduction**

Most researchers have defined procrastination as the lack or absence of self-regulated performance and the behavioral tendency to postpone what is necessary to reach a goal. We procrastinate when we delay beginning or completing an intended course of action (Beswick & Mann, 1994; Ferrari, 1993; Lay & Silverman, 1996). In the academe, procrastination represents a border between the goals and actions of college students because their goal of completing college successfully is held up by their action of procrastinating on studying lessons and doing academic requirements. Not less than 70% of undergraduate students procrastinate, and almost 20% do so regularly (Ellis & Knaus, 1977; O’Brien, 2002; Schouwenburg, 1995), while 50% of them chronically procrastinate (Day, et al., 2000). Chronic procrastination is either getting worse or more people are willing to admit to chronically procrastinating (Steel, 2007). Higher ability students procrastinated more than lower ability students and this behavior become intense as students move on to their academic careers and became more self-regulated (Ferrari, 1991). In most of these studies, researchers have focused on the negative aspects of this behavior and found that procrastination is a maladaptive practice that should be corrected.

Recently, Chu and Choi (2005) provided a different point of view, and showed that there is another facet of academic procrastination other than the negative aspect previous researchers presented. They proposed two different types of procrastinators, passive and active procrastinators. Passive academic procrastinators postpone their tasks because of their inability to make the decision to do tasks in a timely manner. In contrast, active procrastinators make intentional decision to procrastinate, work best under pressure, and are able to complete tasks with satisfactory outcomes. What is interesting with their findings is that active
procrastinators demonstrated similar attitudes, coping styles, and academic performance to those of non-procrastinators.

Other studies have also shown that procrastination presents some benefits too. For instance, Vacha and McBride (1993) found that students who procrastinate were more likely to cram, and that crammers do better than non-crammers by using superior study strategies to attain maximum effectiveness. Sommer (1990) asserted that high ability students maximize the effectiveness of their study time by doing a carefully orchestrated cycle of procrastination and cramming. Brinthaupt and Shin (2001) further investigated the relationship of cramming to maximum efficiency and peak experience. They argued that cramming increases flow because it increases the level of task challenge and demands a higher level of performance from the student.

In an attempt to show the lighter side of academic procrastination, Schraw, Wadkins, and Olafson (2007) used grounded theory to look into this phenomenon. They described the participants in their study as having a wide variety of potentially adaptive characteristics, as well as maladaptive aspects of procrastination. They discovered that adaptive characteristics included cognitive efficiency and peak experience as its dimensions. These findings suggest that procrastination improves efficiency, challenge, and flow. Csikszentmihalyi (1997) defined flow as the state of total involvement in an activity that consumes one’s complete attention. According to Csikszentmihalyi the most important determinant for flow or "optimal experience" is the balance between the challenge of the task or situation and one’s skills. When challenge and skills are matched, flow is more likely to occur. Schraw, et al. (2007) indicated that procrastination ultimately increases the likelihood of achieving a deep state of flow because procrastinators work under pressure for an extended period of time in which all of their resources are focused on one goal.

Similarly, Perry (2008) discussed benefits of procrastination in what he called structured procrastination. This strategy, as he referred to it, converts procrastinators into well-organized human beings and valued for efficiently using their time. He said that structured procrastination is the art of making this bad trait work. The key idea that he offered is that procrastination does not mean absolutely doing nothing. Procrastinators do useful things, as Perry notes, "Procrastinators seldom do absolutely nothing; they do marginally useful things, like gardening or sharpening pencils or making a diagram of how they reorganize their files when they get around to it" (Perry, 2008). The result of this kind of strategy is that in order to avoid that task at the top of our list, we engage in other worthwhile tasks below our priority list.

On the basis of the above-mentioned literature, one thing clear is that there are two types of academic procrastination, the adaptive and the maladaptive procrastination. There is also no definite instrument to date that could clearly measure and distinguish the adaptive and maladaptive procrastination. Although researchers in the field of social and personality psychology have recognized this new development on procrastination (e.g. Alexander & Onwuegbuzie, 2007; Bui, 2007; Howell & Watson, 2007; Hu,
Huhmann, & Hyman, 2007; Vacha & McBride, 1993), few studies have expanded on this idea and that no researcher in the field of educational psychology has taken serious steps in making a reliable and valid adaptive/maladaptive academic procrastination instrument.

In an effort to address these dilatory behaviors especially the emerging positive dimension of this behavior, Morales (2010) developed a 65-item Academic Procrastination Scale (APS). The APS was developed to measure three dimensions of academic procrastination (i.e., structured procrastination, unstructured procrastination, and non-procrastination).

**Objectives of the Study**

The main objective of the study was to validate the Academic Procrastination Scale that assesses each of the three academic procrastination constructs. It determined if indeed an internally consistent measure of structured procrastination could be used in a classroom context. Of particular interest was whether the three procrastination constructs could be validated as statistically independent using confirmatory factor analysis (CFA). It also sought to determine if indeed the items for each construct would fall under each category. It is predicted that CFA of the three academic procrastination dimensions would yield acceptable goodness of fit indices. This study also examined the means and intercorrelations among the academic procrastination variables with specific interest in structured procrastination in relation to the other constructs.

**Methods**

**Participants**

The sample was drawn from five colleges in a state university in Region 8 in central Philippines. A total of 1200 students enrolled for the 2nd semester of school year 2009-2010 were given questionnaires but only 1153 were used in actual analysis because some questionnaires were not retrieved or others have either no response or missing necessary information. The students came from the Colleges of Education, Science, Nursing, Engineering, and Business Administration.

**Instrument**

The newly developed 65-item academic procrastination scale in Study 1 features three dimensions of academic procrastination. The first dimension, structured procrastination (α=.93), measures the positive side of procrastination. Sample items include “I always manage to finish report papers even if I started it few hours before deadline” and “I intentionally put off work to maximize my motivation.” Unstructured procrastination (α = .92) is the second dimension measuring procrastination related to fear of failure, task aversiveness and laziness. Items include “I don’t think I have
enough knowledge to write a school paper" and “I set aside reading lesson for a test when my friends drop by at our house.” The third dimension, non-procrastination (α = .78) clusters items related to the non-practice of procrastination. Items such as “I usually accomplish all the things I plan to do in a day” and “I tend to finish tasks well ahead of deadlines” are included. Participants answered each item using a 6-point scale that ranges from 1 (strongly disagree) to 6 (strongly agree).

**Data Analysis**

A two-step analysis was employed in this study. First, descriptive statistics was used to present the basic facts of all the variables involved. These preliminary analyses examined whether basic characteristics of the current data set (i.e. means, standard deviations, percentages, skewness, kurtosis) are acceptable for further analyses. It also included the examination of reliability coefficients and relationships of factor structures of the measure. The purpose of examining estimates of internal consistency from the sample was to determine if the measures that were used have acceptable reliability levels or reliability estimates. Bivariate relation between the factors of procrastination processes was conducted to determine how each variable associate itself with other variables.

Second, Confirmatory factor analysis (CFA) for the measurement model (i.e., Academic Procrastination) was conducted. CFA was investigated using Structural Equation Modelling (SEM) that uses maximum likelihood estimation (MLE). This was followed by assessment of model fit to determine the degree to which the measurement model fits the data (Joreskog & Sorbom, 1989). In evaluating the fit of the models, recommendations by Schermelleh-Engel, Moosbrugger, & Müller (2003) were followed. These recommendations state that for an acceptable model fit, the ratio \( \chi^2/df \) should be less than or equal to 5, the Root Mean Square Error of Approximation (RMSEA) should be less than or equal to .08, the standardized root-mean-square (SRMR) should be less than .05, the Tuker-Lewis Index (TLI) should be greater than or equal to .95, and the Comparative Fit Index (CFI) should also be greater than or equal to .90 (Schermelleh-Engel, et al., 2003). The RMSEA, SRMR, TLI, and CFI were chosen because they were found as being less affected by the size of the sample when compared to the Normative Fit Index (NFI), the Goodness-of-Fit Index (GFI), and the Adjusted Goodness-of-Fit Index (AGFI) (Schermelleh-Engel, et al., 2003).

**Procedure**

Participants were asked to fill out an informed consent stating that they were volunteers and could end answering the questionnaire at any time. The consent form also had the researcher’s name and email address in case there were any future questions. All participants were given brief verbal instructions by the researcher or the faculty in-charge. After administering the questionnaires, data were encoded and cleaned for
errors (e. g., typographical, missing personal information, incomplete entries). Items comprising each of the 3 variables were taken as is in the analysis to represent score for each variable. Negatively keyed items were scored in reverse.

**Results and Discussion**

**Descriptive Statistics**

The means, standard deviations, skewness, and kurtosis values for the 65 items on the APS are reported in Table 1. Item means indicated a moderate ceiling effect. The skewness and kurtosis values for all APS items were within acceptable range of ± 1.96 (George & Mallery, 2001), suggesting no concern about deviation from normality.

**Table 1**

*Means, Standard Deviations, Skewness, and Kurtosis for the Academic Procrastination scale items*

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Structured Procrastination</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP2</td>
<td>4.26</td>
<td>1.39</td>
<td>-0.67</td>
<td>-0.36</td>
</tr>
<tr>
<td>AP4</td>
<td>3.68</td>
<td>1.42</td>
<td>-0.12</td>
<td>-0.81</td>
</tr>
<tr>
<td>AP5</td>
<td>4.00</td>
<td>1.28</td>
<td>-0.40</td>
<td>-0.33</td>
</tr>
<tr>
<td>AP8</td>
<td>3.57</td>
<td>1.48</td>
<td>-0.04</td>
<td>-0.88</td>
</tr>
<tr>
<td>AP9</td>
<td>3.59</td>
<td>1.36</td>
<td>-0.26</td>
<td>-0.65</td>
</tr>
<tr>
<td>AP12</td>
<td>3.76</td>
<td>1.41</td>
<td>-0.27</td>
<td>-0.73</td>
</tr>
<tr>
<td>AP13</td>
<td>3.48</td>
<td>1.33</td>
<td>-0.11</td>
<td>-0.66</td>
</tr>
<tr>
<td>AP14</td>
<td>3.31</td>
<td>1.33</td>
<td>-0.05</td>
<td>-0.79</td>
</tr>
<tr>
<td>AP17</td>
<td>2.83</td>
<td>1.26</td>
<td>0.31</td>
<td>-0.62</td>
</tr>
<tr>
<td>AP18</td>
<td>3.28</td>
<td>1.19</td>
<td>0.00</td>
<td>-0.55</td>
</tr>
<tr>
<td>AP20</td>
<td>3.54</td>
<td>1.27</td>
<td>-0.12</td>
<td>-0.55</td>
</tr>
<tr>
<td>AP21</td>
<td>3.64</td>
<td>1.30</td>
<td>-0.20</td>
<td>-0.65</td>
</tr>
<tr>
<td>AP22</td>
<td>3.55</td>
<td>1.29</td>
<td>-0.15</td>
<td>-0.41</td>
</tr>
<tr>
<td>AP23</td>
<td>3.59</td>
<td>1.19</td>
<td>-0.19</td>
<td>-0.56</td>
</tr>
<tr>
<td>AP24</td>
<td>3.48</td>
<td>1.37</td>
<td>-0.23</td>
<td>-0.71</td>
</tr>
<tr>
<td>AP25</td>
<td>2.84</td>
<td>1.33</td>
<td>0.40</td>
<td>-0.50</td>
</tr>
<tr>
<td>AP26</td>
<td>3.02</td>
<td>1.29</td>
<td>0.18</td>
<td>-0.61</td>
</tr>
<tr>
<td>AP31</td>
<td>3.08</td>
<td>1.24</td>
<td>-0.15</td>
<td>-0.76</td>
</tr>
<tr>
<td>AP32</td>
<td>2.88</td>
<td>1.16</td>
<td>0.12</td>
<td>-0.60</td>
</tr>
<tr>
<td>AP39</td>
<td>3.95</td>
<td>1.32</td>
<td>-0.48</td>
<td>-0.37</td>
</tr>
<tr>
<td>AP40</td>
<td>3.06</td>
<td>1.29</td>
<td>0.20</td>
<td>-0.58</td>
</tr>
<tr>
<td>AP43</td>
<td>3.12</td>
<td>1.35</td>
<td>0.07</td>
<td>-0.79</td>
</tr>
<tr>
<td>AP44</td>
<td>3.17</td>
<td>1.25</td>
<td>0.10</td>
<td>-0.65</td>
</tr>
<tr>
<td>AP47</td>
<td>3.27</td>
<td>1.33</td>
<td>-0.10</td>
<td>-0.83</td>
</tr>
<tr>
<td>AP48</td>
<td>3.25</td>
<td>1.28</td>
<td>0.00</td>
<td>-0.63</td>
</tr>
</tbody>
</table>
CFA was conducted on the academic procrastination model using AMOS 7 (Arbuckle, 1997). The analyses were conducted on covariance matrices, and the solutions were generated on the basis of maximum-likelihood estimation. The model was treated as observed variables by taking the average from its respective items. Following Schermelleh-Engel, et al. (2003), both absolute (e.g., chi square) and incremental (e.g., Comparative Fit Index [CFI]) fit indices were used to evaluate the fit of the models to the data.

Confirmatory Factory Analysis examined the academic procrastination model, which designated that the items for each type of procrastination load on their respective latent variables. The results from this analysis supported the hypothesized model, as each fit statistic met the acceptable criteria for a good fitting model: $\chi^2 / df = 4.33$; root-mean-square error of approximation (RMSEA) = .047; SRMR = .0499; Tucker–Lewis Index (TLI) = .93; CFI = .90. Factor loadings for this model presented show that value for each item ranged from .46 to .93. The CFA data clearly indicate that the three academic procrastination measures represent empirically separable and internally consistent variables. These results further suggest that the three-factor model represents a good fit to
the responses of the participants to the 65-APS items. Table 2 presents standardized factor loadings for this three-factor model. The factor loadings ranged from .45 to .75 for the first factor, from .43 to .90 for the second factor, and from .46 to .98 for the third factor. The factors were positively related to each other with correlation coefficients ranging from .34 to .88. All factor loadings were statistically significant, \( p < .05 \).

**Table 2**  
*Standardized Factor Loadings for the Three-Factor Model of the APS Items*

<table>
<thead>
<tr>
<th>Items</th>
<th>Structured</th>
<th>Factor Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>AP2</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>AP4</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>AP5</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>AP8</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>AP9</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>AP12</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>AP13</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>AP14</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>AP17</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>AP18</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>AP20</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>AP21</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>AP22</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>AP23</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>AP24</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>AP25</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>AP26</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>AP31</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>AP32</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>AP39</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>AP40</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>AP43</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>AP44</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>AP47</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>AP48</td>
<td>0.80</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unstructured</th>
<th></th>
<th>0.50</th>
<th>0.63</th>
<th>0.51</th>
<th>0.78</th>
<th>0.52</th>
<th>0.66</th>
<th>0.84</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bivariate Relations among the Academic Procrastination

Table 3 shows that the zero-order correlations among the measures indicate that structured procrastination was positively associated with unstructured procrastination \((r = .26, p < .01)\). Both structured \((r = .11, n.s.)\) and unstructured \((r = -.05, n.s.)\) procrastination were not associated with non-procrastination. It should be noted that similar relationships were observed during the development phase where structured procrastination was positively correlated with unstructured procrastination but both types of procrastination were not significantly associated with non-procrastination. The low coefficient yet highly significant relationship between structured and unstructured procrastination shows the discriminant validity of the two constructs. They are conceptually different yet theoretically similar constructs. Both are dilatory behaviors but are done with different intentions and outcomes.

<table>
<thead>
<tr>
<th></th>
<th>Structured</th>
<th>Unstructured</th>
<th>Non-procrastination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstructured</td>
<td>.26**</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Non-procrastination</td>
<td>.11 n.s.</td>
<td>-.05 n.s.</td>
<td>-</td>
</tr>
</tbody>
</table>
Means, Standard Deviation, and Reliabilities of the factors

Internal consistency reliabilities for the scale scores were .91 for structured procrastination, .89 for unstructured procrastination and .76 for non-procrastination.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>M</th>
<th>SD</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured Procrastination</td>
<td>1-6</td>
<td>3.21</td>
<td>0.71</td>
<td>.91</td>
</tr>
<tr>
<td>Unstructured Procrastination</td>
<td>1-6</td>
<td>3.26</td>
<td>0.73</td>
<td>.89</td>
</tr>
<tr>
<td>Non-procrastination</td>
<td>1-6</td>
<td>3.93</td>
<td>0.75</td>
<td>.76</td>
</tr>
</tbody>
</table>

Confirmatory factor analysis documented that each of the procrastination tendencies represents different constructs. The factor analytic technique also indicated that the academic procrastination framework provided a good fit to the data. Examination of the structured procrastination score revealed that this dilatory behavior is clearly operative in the undergraduate classroom (the means were close above the scale midpoint, and the full range of scores was used). In addition, structured procrastination was consistently correlated with the negative procrastination construct with which they shared a theoretically similar constructs that bond positively. On the other hand, both structured and unstructured procrastination evidenced no association with non-procrastination both in this study and in the scale development phase.

In general, the relations among variables involved in this study largely support the hypotheses forwarded. This demonstrates the generalizability of the newly developed academic procrastination framework among Filipino college students. The study found structured procrastination construct to be valid and reliable measure different from the negative procrastination construct.

Generally, this study highlights the importance of knowing that there exists a procrastination behavior with positive outcome, different from traditional view of procrastination with negative consequence. This finding provides important implication for teachers. They should be aware that students' postponement of tasks could be their strategy of making the task interesting and that teachers should not jump into conclusion and make a negative impression about it. It is hoped that the academic procrastination framework established in this study will serve as additional theoretical and empirical tool in addressing this issue and the
many other important issues that await attention in the academic procrastination literature.

References


**About the Author**

Dr. Rommel Morales is presently teaching statistics and assessment at the college of education, University of Eastern Philippines, Northern Samar. He finished his PhD in Educational Psychology major in quantitative analysis at De La Salle University, Manila.

For more inquiries, the author can be reached at: University of Eastern Philippines – Catarman, Northern Samar

E-mail: mel.mrls@yahoo.com