CONTEMPORARY EDUCATIONAL PSYCHOLOGY 19, 460–475 (1994)

Assessing Metacognitive Awareness

GREGORY SCHRAW AND RAYNE SPERLING DENNISON
Department of Educational Psychology, University of Nebraska at Lincoln

We constructed a 52-item inventory to measure adults' metacognitive awareness. Items were classified into eight subcomponents subsumed under two broader categories, knowledge of cognition and regulation of cognition. Two experiments supported the two-factor model. Factors were reliable (i.e, α = .90) and intercorrelated (r = .54). Experiment 2 reported the knowledge of cognition factor was related to pre-test judgments of monitoring ability and performance on a reading comprehension test, but was unrelated to monitoring accuracy. Implications for educational assessment and future research were discussed.

Metacognition refers to the ability to reflect upon, understand, and control one's learning. Previous accounts of metacognition have distinguished between two major components, including knowledge about cognition and regulation of cognition (Brown, 1987; Flavell, 1987; Jacobs & Paris, 1987). *Knowledge about cognition* includes three subprocesses that facilitate the reflective aspect of metacognition: declarative knowledge (i.e., knowledge about self and about strategies), procedural knowledge (i.e., knowledge about how to use strategies), and conditional knowledge (i.e., knowledge about when and why to use strategies). *Regulation of cognition* includes a number of subprocesses that facilitate the control aspect of learning. Five component skills of regulation have been discussed extensively, including planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation (Artzt & Armour-Thomas, 1992; Baker, 1989).

Recent research indicates that metacognitively aware learners are more strategic and perform better than unaware learners (Garner & Alexander, 1989; Pressley & Ghatala, 1990). One explanation is that metacognitive awareness allows individuals to plan, sequence, and monitor their learning in a way that directly improves performance. Swanson (1990), for example, found that metacognitively aware sixth graders used more strategies and performed better than unaware students when solving fluid combination and pendulum problems. Differences in strategy use and

We thank Dorie Munson, Barbara S. Plake, and Dan Robinson for their helpful comments on items appearing in an earlier draft of the MAI. Address reprint requests to Gregory Schraw, Department of Educational Psychology, 1313 Seaton Hall, University of Nebraska, Lincoln, NE 68588-0641.

0361-476X/94 $6.00

performance were related to differences in metacognitive awareness rather than differences in intellectual aptitude. This finding suggested that metacognitive knowledge plays a compensatory role in cognitive performance by improving strategy use. Artzt and Armour-Thomas (1992) also found that metacognitive awareness of mathematical problem solving strategies among grade school children was associated with the increased use of problem-solving heuristics and higher-level group responses to problems. In a study comparing adults, Corkill and Koshida (1993) reported that metacognitive awareness measured during on-line interviews was related to monitoring accuracy.

In addition, a number of unexpected findings have occurred that are at odds with intuitive assumptions about metacognition. One is that metacognitive awareness appears to be independent of intellectual ability (Swanson, 1990) and academic achievement (Pressley & Ghatala, 1990). A second is that regulatory skills such as monitoring may be independent or even negatively related to domain knowledge (Glenberg & Epstein, 1987). A third is that metacognitive monitoring appears to be independent of ease of comprehension (Leonsario & Nelson, 1990). These findings are important because they reveal that metacognition is separable from other cognitive constraints on learning such as aptitude and domain knowledge and cannot be predicted entirely, or even with a moderate degree of accuracy, on their basis.

In light of these findings, one of the most difficult problems facing researchers and practitioners is identifying metacognitively aware learners quickly and reliably. Virtually all previous investigations have identified highly metacognitive learners using some form of on-line experimental testing such as identifying anomalous text segments (Markman, 1977), calibrating one's comprehension (Pressley & Ghatala, 1990), or through extensive verbal interviews conducted prior to problem solving (Artzt & Armour-Thomas, 1992; Corkill & Koshida, 1993; Swanson, 1990). These procedures are prohibitive in most applied settings due to the amount of time and effort necessary to administer them. For example, Swanson (1990) asked students to respond individually to a series of 17 questions intended to assess metacognitive awareness. While apparently reliable, Swanson's method no doubt places excessive demands on researchers and classroom teachers interested in metacognitive ability from a diagnostic perspective.

The purpose of the present research was to generate and test an easily administered metacognitive inventory suitable for adolescents and adults. Three aspects of this question were of particular importance to us. First, we investigated whether current conceptualizations of metacognition that postulate two main processes appear to be valid. To test this hypothesis, we developed a 52-item self-report instrument that included multiple items within each of the eight component processes subsumed under knowledge and regulation of cognition. We conducted a set of *unrestricted* analyses in which the 52 items created multiple, distinct factors. These solutions were used to examine the match between the eight theoretical subcomponents subsumed under knowledge and regulation of cognition and observed loading patterns. We also conducted a set of *restricted* analyses in which the 52 items were forced onto two factors. These solutions were used to examine the match between knowledge and regulation of cognition processes and observed loading patterns.

A second issue addressed the statistical relationship between knowledge and regulation of cognition. One key assumption of all theories of metacognition is that knowledge and regulation of cognition are mutually correlated and compensatory (Baker, 1989; Brown, 1987; Flavell, 1987; Jacobs & Paris, 1987). This relationship was addressed by comparing a number of factor-analytic solutions of our instrument. Oblique solutions (i.e., correlated factors) were expected to be superior to orthogonal solutions (i.e., uncorrelated factors).

Third, Experiment 2 investigated the convergent validity of the instrument by comparing the relationship between knowledge and regulation of cognition with measures of pre-test monitoring ability, test performance, and the ability to monitor one's test performance accurately. We predicted that estimates of monitoring ability, normatively high test performance, and accurate monitoring would be associated with higher scores on the 52-item instrument.

EXPERIMENT 1

Method

Participants. One hundred and ninety-seven undergraduates (112 females, 85 males) enrolled in an introductory educational psychology course at a large midwestern university participated in the study as part of their normal course assignment.

Materials. The materials consisted of a 52-item self-report instrument (henceforth referred to as the Metacognitive Awareness Inventory, or MAI) developed for the purposes of this study by the authors. The development of the instrument was exploratory in nature in that we intended to compare two- and eight-factor solutions. The two-factor solution corresponded to knowledge and regulation of cognition. The eight-factor solution corresponded to the eight subcomponents subsumed under knowledge and regulation of cognition.

An initial pool of 120 items were written, including at least eight items in each of the eight categories described above. Items were piloted on a group of college undergraduates (N = 70), revised, and eliminated where appropriate. Items with extreme mean scores (i.e., near either end of the bi-polar scale) were dropped, as were scores with an unusually high degree of variability. In addition, some highly intercorrelated items were dropped such that only one of these items remained on the scale. The final version of the instrument consisted of 52 items distributed across the eight scales with at least four items per scale (see Appendix A).

The letters enclosed within parentheses in Appendix A indicate the theoretical scale on which the item was included. The eight scales are as follows: (1) declarative knowledge, (2) procedural knowledge, (3) conditional knowledge, (4) planning, (5) information management strategies, (6) monitoring, (7) debugging strategies, and (8) evaluation of learning. Operational definitions for each category appear in Appendix B.

Ratings for each item were made on a 100-mm, bi-polar scale adapted from the multidimensional scaling literature. The right end of each scale indicated the statement was false about the individual; the left end indicated the statement was true. Continuous scales were used in place of Likert scales for several reasons. First, continuous scales provide a better approximation to interval data compared to Likert scales. Second, scoring responses on a 100-point scale compared to a 5- or 7-point scale was expected to increase response variation which in turn should enhance the reliability of the instrument. Subjects recorded their responses by drawing a slash across the rating scale at a point that best corresponded to how true or false the statement was about them.

Inventories also included brief cover instructions informing individuals that the purpose of the instrument was to assess knowledge about learning. A brief description and an example of the rating scale were included in these instructions. Individuals were told that responses were anonymous and would not affect their class grade in any way and that they should answer each item as carefully and truthfully as possible.

Procedure. Individuals were given the instrument during their regularly scheduled class period and asked to complete and return it at their next class meeting (2 days). Pilot testing revealed that the average completion time was approximately 10 min. Inventories that were not returned the next class period were discarded.

Results

The instrument was evaluated using a number of unrestricted and restricted (i.e., forced solution) factor analyses. Each of these analyses is described below.

Unrestricted factor analyses. Separate orthogonal (i.e., uncorrelated factors) and oblique (i.e., correlated factors) solutions were performed to assess the match between the eight hypothesized subscales and observed factors. Both analyses produced six-factor solutions with eigenvalues greater than one that recovered 78% of the sample variance. Factor loadings for each for the orthogonal and oblique solutions were virtually identical. However, the oblique solution indicated that each pair of factors was correlated in excess of r = .30, whereas the orthogonal rotation forced factors to be mutually uncorrelated.

An inspection of the oblique and orthogonal six-factor solutions revealed that neither corresponded precisely to the eight hypothesized subscales described above. Factors 1 and 2 included most of the items in the knowledge of cognition category (i.e., declarative, procedural, and conditional knowledge), whereas Factors 3 through 6 included most of the items in the regulation of knowledge category. In addition, coefficient α for most of the six factors was below our desired criterion value of .80: .81, .74, .71, .66, .65, and .59 respectively.

Overall, the unrestricted factor solutions did not lead to a parsimonious interpretation of the data, nor to highly reliable factors. We therefore concluded that the MAI does not measure a number of distinct subcomponents of metacognition and, in particular, does not correspond closely to the eight subcomponents described above. For this reason, we performed a forced, two-factor solution on the data to compare whether these factors corresponded to the knowledge of cognition and regulation of cognition dimensions proposed by Brown (1987) and Jacobs and Paris (1987).

Restricted factor analyses. Both oblique and orthogonal two-factor solutions were performed on the data. Both solutions resulted in virtually identical solutions with the exception that the two factors were highly correlated in the oblique solution, $r = .54$. As expected, items included in the knowledge of cognition category (i.e., declarative, procedural, and conditional knowledge) loaded on the first factor while items included in the regulation of cognition category loaded on the second factor. A small percentage of items had loadings greater than .30 on both factors (i.e., items 27, 34, 35, 40, 43, and 44) while two items failed to load on either factor (i.e., items 4 and 48). Coefficient α for items loading on each factor reached .91, indicating a high degree of internal consistency. Coefficient α for the entire instrument reached .95.

Table 1 presents loadings from the oblique solution for each item (i.e., loadings from the structure matrix). Twenty-five items loaded unambiguously on Factor 1 while 19 items loaded on Factor 2. The two factors accounted for 65% of the sample variance. Items with high loadings on Factor 1 were those in the knowledge of cognition category. Items with high loadings on Factor 2 ranged across each of the five subscales included in the regulation of cognition category.

Discussion

The results of the exploratory analyses described above suggest two conclusions. First, there is little evidence in support of the eight subcomponents of metacognition described earlier. Six rather than eight factors were obtained, none of which bore a close resemblance to any of the eight predicted factors. Second, the forced, two-factor solution yielded a solution in close accord with theoretical predictions. Items originally classified under knowledge of cognition loaded on Factor 1; items classified under regulation of cognition loaded under Factor 2. Both factors were stable and reliable. In addition, few items loaded highly on both factors. Of the six items that did, all had moderately low loadings on both factors (i.e., less than .45). In addition, five of the six items came from different hypothesized subscales, suggesting that there was no systematic cross loading among items.

From a psychometric standpoint, the two-factor solution indicated that the MAI measured two kinds of metacognitive knowledge reliably.

TABLE 1
LOADINGS FOR THE TWO-FACTOR OBLIQUE SOLUTIONS IN EXPERIMENTS 1 AND 2

Item	Factor 1	Factor 2
32	.70 (.69)	.00 (.00)
26	.66 (.52)	.00 (.00)
5	.65 (.43)	.00 (.00)
30	.59 (.59)	.00 (.00)
46	.57 (.30)	.00 (.00)
10	.56 (.72)	.00 (.00)
13	.55 (.66)	.00 (.00)
3	.55 (.54)	.00 (.00)
29	.54 (.35)	.00 (.00)
16	.53 (.51)	.00 (.00)
15	.53 (.41)	.00 (.00)
9	.51 (.36)	.00 (.32)*
20	.51 (.37)	.00 (.00)
52	.48 (.00)*	.00 (.00)
33	.48 (.57)	.00 (.00)
12	.46 (.65)	.00 (.00)
18	.43 (.00)*	.00 (.39)*
17	.42 (.57)	.00 (.00)
7	.41 (.41)	.00 (.00)
51	.40 (.00)*	.00 (.51)*
39	.38 (.00)*	.00 (.00)
45	.38 (.32)	.00 (.00)
25	.34 (.00)*	.00 (.00)
31	.34 (.30)	.00 (.00)
42	.31 (.38)	.00 (.00)
22	.00 (.00)	.70 (.52)
36	.00 (.00)	.70 (.67)
8	.00 (.00)	.68 (.55)
50	.00 (.00)	.65 (.00)*
38	.00 (.00)	.62 (.64)
1	.00 (.00)	.62 (.45)
23	.00 (.00)	.60 (.63)
6	.00 (.00)	.59 (.44)
49	.00 (.00)	.55 (.55)
21	.00 (.00)	.55 (.00)*
24	.00 (.00)	.52 (.34)
28	.00 (.00)	.50 (.31)
11	.00 (.00)	.46 (.43)
14	.00 (.00)	.46 (.40)
19	.00 (.00)	.44 (.36)
47	.00 (.30)*	.39 (.00)*
37	.00 (.00)	.38 (.34)
2	.00 (.00)	.36 (.58)
41	.00 (.00)	.32 (.30)
27	.42 (.35)	.37 (.00)*
43	.37 (.41)	.37 (.00)*

TABLE 1—*Continued*

Item	Factor 1	Factor 2
44	.34 (.00)*	.41 (.49)
34	.34 (.40)	.36 (.32)
40	.34 (.41)	.30 (.00)*
35	.31 (.36)	.43 (.41)
4	.00 (.00)	.00 (.00)
48	.00 (.00)	.00 (.00)

Note. Unenclosed factor loadings are from Experiment 1; enclosed loadings are from Experiment 2.

*Indicates differences in loadings across Experiments 1 and 2.

Knowledge of cognition corresponded to what students know about themselves, strategies, and the conditions under which strategies are most useful. Regulation of cognition corresponded to knowledge about the ways that students plan, implement strategies, monitor, correct comprehension errors, and evaluate their learning. These factors also were strongly intercorrelated, suggesting that knowledge and regulation may work in unison to help students self-regulate.

EXPERIMENT 2

The purpose of Experiment 2 was to validate the MAI using empirically derived measures of metacognitive knowledge, test performance, and metacognitive regulation. The metacognitive knowledge variable consisted of a pre-test judgment of monitoring ability. The test performance component included scores from four reading comprehension tests. The metacognitive regulation component consisted of confidence judgments about test performance.

If the MAI measures metacognitive awareness among adults, it should be correlated with test performance and empirical measures of metacognition. Specifically, higher scores on the MAI, indicating greater metacognitive awareness, should correspond to higher test performance, a greater awareness of one's own monitoring skills, and accurate monitoring of one's test performance.

Method

Participants. One hundred and ten undergraduates (69 females, 41 males) enrolled in an introductory educational psychology class participated as part of their course requirement.

Materials. Individuals completed the MAI as described in Experiment 1. The instrument included brief cover instructions, 52 items, and a 100-mm rating scale following each item. Individuals also completed five reading comprehension subtests from the Nelson-Denny reading comprehension test (Form E) (Brown, Bennett, & Hanna, 1981). The first of these tests was used as practice. The remaining four tests included 16 four-option multiple choice test questions covering four short expository passages.

Individuals also rated their confidence for each of the 16 test questions using a 100-mm scale appearing after each question. The left end of each scale was labeled *0% Confidence*, whereas the right end was labeled *100% Confidence*. Individuals were instructed to draw a slash on the line at the point that best corresponded to their perceived confidence.

In addition, each person rated his or her perceived monitoring ability prior to completing the reading comprehension test. The pre-test ability scale consisted of a bipolar, 100-mm rating scale identical to those described above. The left end of the scale was labeled *Poor Monitoring Ability*, whereas the right end was labeled *Excellent Monitoring Ability*. The purpose of this scale was to assess individuals' knowledge about their monitoring competence. Individuals rated their monitoring ability by drawing a slash on the scale at the point that best corresponded to their perceived ability to monitor their comprehension. Instructions for the rating task were as follows:

> We want you to rate how well you can monitor the accuracy of your performance on multiple-choice tests based on short, technical reading passages. If you generally know how well you do on reading comprehension tests such as these, draw a slash near the end of the scale labeled *Excellent Monitoring Ability*. If you generally are unable to monitor your test performance accurately, draw a slash toward the end labeled *Poor Monitoring Ability*. Please make your rating on any part of the line that best corresponds to your perceived monitoring ability.

Procedure. Individuals participated in groups of 15 to 25 people. The entire testing session lasted approximately 45 min. All participants received experimental materials in the same order. The order of completion was as follows: Individuals first were asked to complete the MAI. General instructions were given next regarding the reading comprehension phase of the study. Individuals next rated their monitoring ability, completed the practice reading passage, then completed the four reading comprehension tests. The format for each test was identical: Individuals read the story, turned to the next page in their test booklet to complete the test, then rated how much confidence they had in each response. Participants were not allowed to look back at text passages once they began the test. There were no time limits on any phase of the experiment.

Results

Four analyses were conducted to assess the predictive validity of the MAI. The first replicated the oblique two factor structure reported in Experiment 1. The remaining analyses compared scores on the MAI with pre-test judgments of monitoring ability, test performance, and monitoring accuracy. Monitoring accuracy was computed by taking the difference between each person's average confidence on the 16 test items and their actual test score expressed as a proportion. Both test performance accuracy and confidence in one's performance ranged from zero to 100. Difference scores in excess of zero corresponded to overconfidence. Scores below zero corresponded to underconfidence. Scores equal to zero corresponded to perfect accuracy. Means and standard deviations across the entire sample are presented in Table 2.[1] All tests were conducted at the $p < .05$ level of significance.

[1] To facilitate data interpretation, individuals were assigned to three mutually exclusive groups based on pre-test judgments, test performance, and monitoring accuracy. It is important to note that the use of multiple regression analyses for continuous variables would

TABLE 2
MEANS AND STANDARD DEVIATIONS FOR VARIABLES USED IN EXPERIMENT 2

	Mean	*SD*
Pre-test ratings	67.27	16.33
Test performance	65.95	17.64
Confidence ratings	79.30	13.02
Monitoring accuracy	13.35	17.86
Knowledge of cognition	77.02	12.78
Regulation of cognition	67.45	12.87

Note. Pre-test, confidence, knowledge of cognition, and regulation of cognition ratings were made on a 0–100 scale. Test performance is expressed as the proportion correct. Monitoring accuracy equals the difference between confidence and performance scores.

Factor replication. Oblique and orthogonal two-factor solutions were performed on the data. These solutions corresponded quite closely to Experiment 1 (see Table 1 for a comparison). The two factors corresponded to knowledge and regulation of cognition and accounted for 58% of the sample variance. Coefficient α for Factor 1 (i.e., knowledge of cognition) and Factor 2 (i.e., regulation of cognition) reached .88 and .88, respectively. Coefficient α for the entire instrument reached .93. The correlation between the two factors reached $r = .45$.

A comparison of the loadings in Table 1 indicated that both solutions yielded similar loading structures with two exceptions. One was that items 18 and 51 loaded on different factors in Experiments 1 and 2. Another is that several items (e.g., 21, 50, and 52) failed to load on either factor even though each had a meaningful loading in Experiment 1. Notwithstanding these minor differences, Factors 1 and 2 led to the same interpretation in both experiments. Factor 1 corresponded to knowledge of cognition; Factor 2 corresponded to regulation of cognition.

We next constructed two scores based on the oblique solution that could be used as dependent variables in subsequent analyses. Items with a loading in excess of .30 on Factor 1 and less than .30 on Factor 2 were used to create a mean score for *knowledge of cognition*. This score was based on 19 items. A similar process was used to create a mean *regulation of cognition* score that included 22 items.

The relationship between pre-test judgments and the MAI. Individuals

lead to identical results with the exception that each analysis would contain several additional degrees of freedom. The small loss of degrees of freedom given our sample size was not expected to affect the power of these tests. In contrast, partitioning independent variables into mutually exclusive groups was expected to make comparisons among extreme groups considerably more straightforward.

were partitioned into three groups on the basis of pre-test judgments of monitoring ability. Trichotomies were formed to facilitate comparisons between extreme groups. Group 1 (N = 38) scored below 65 on the 100-point rating scale. Group 2 (N = 46) scored between 66 and 80 on this scale. Group 3 (N = 26) scored above 80 on the scale. In addition, correlations among variables used in Experiment 2 are shown in Table 3.

Pre-test judgments were used as a measure of metacognitive knowledge about one's monitoring skills. We predicted that judgments of high monitoring accuracy would be associated with high scores on the MAI's knowledge of cognition factor. A one-way multivariate analysis of variance (MANOVA) using knowledge and regulation of cognition as dependent variables did not reach significance, *approximate* $F(6,210) = 1.89$, $MS_e = .476$. However, several one-way analyses of variance (ANOVA) revealed that knowledge of cognition scores differed significantly across the three groups, $F(2,107) = 4.45$, $MS_e = 681.91$, whereas regulation of cognition scores did not. A Tukey's Honestly Significant Difference test (HSD) revealed that Group 3 (M = 81.37, SD = 12.10) scored significantly higher than Group 1 (M = 72.86, SD = 15.12). Group 2 (M = 78.07, SD = 11.34) also scored higher than Group 1, but did not differ from Group 3.

Additional ANOVAs found that the three groups differed with respect to test performance, $F(2,107) = 6.89$, $MS_e = .192$, but did not differ with respect to monitoring accuracy. A Tukey's HSD test revealed that Group 3 (M = .72, SD = .16) answered a greater proportion of test questions correctly than Group 2 (M = .62, SD = .16) and Group 1 (M = .56, SD = .18). Group 2 also outperformed Group 1. An examination of monitoring scores revealed the expected trends, although group differences were not significant. For example, Group 3's scores (M = .10, SD = .19) were closer to perfect accuracy (i.e., .00) than Group 2 (M = .17, SD = .16) and Group 1 (M = .16, SD = .18). Our failure to find significant statistical differences in this case may be due to the high amount of variability within groups rather than lack of differences between groups.

TABLE 3
ZERO-ORDER CORRELATIONS BETWEEN VARIABLES USED IN EXPERIMENT 2

	1	2	3	4	5	6
Pretest	—	.31**	.12	.26**	.51**	−.19*
KOC		—	.49**	.20*	.23*	.09
ROC			—	.06	.21*	.09
Performance				—	.32**	−.79**
Confidence					—	.32**
Accuracy						—

*$p < .05$; **$p < .01$.

The relationship between test performance and the MAI. Individuals were partitioned into three groups on the basis of test performance. Average performance on the test was 66% correct across the entire sample. Group 1 (i.e., poor performers, $N = 43$) scored below the 60th percentile on the test. Group 2 (i.e., average performers, $N = 47$) scored between the 61st and 75th percentiles. Group 3 (i.e., high performers, $N = 20$) scored above the 75th percentile.

A one-way MANOVA using knowledge and regulation of cognition scores as dependent variables reached significance, *approximate* $F(6,208) = 2.31$, $MS_e = .380$. Subsequent univariate tests revealed a significant effect for knowledge of cognition, $F(2,107) = 6.22$, $MS_e = 926.86$, but failed to report an effect for regulation of cognition. A comparison of individual means using Tukey's HSD procedure revealed that Group 3 reported significantly higher knowledge of cognition scores ($M = 85.15$, $SD = 9.05$) than either Group 2 ($M = 77.14$, $SD = 10.55$) or Group 1 ($M = 73.30$, $SD = 14.79$), which did not differ from each other.

An additional ANOVA found a significant main effect for monitoring accuracy across the three levels of performance, $F(2,107) = 40.69$, $MS_e = .746$. A Tukey's HSD test revealed that Group 3 ($M = -.04$, $SD = .09$) was significantly more accurate (i.e., less overconfident) than Group 2 ($M = .10$, $SD = .13$) or Group 1 ($M = .28$, $SD = .16$). Group 2 also was more accurate than Group 1.

The relationship between monitoring accuracy and the MAI. Individuals were partitioned into three groups on the basis of monitoring accuracy. Group 3 (i.e., accurate monitors, $N = 35$) received accuracy scores less than .05. These individuals tended to be highly accurate or slightly underconfident. Group 2 (i.e., slightly overconfident monitors, $N = 34$) received accuracy scores between .06 and .20. Group 1 (i.e., highly overconfident monitors, $N = 41$) received scores over .20.

Neither a MANOVA or separate ANOVAs revealed significant differences between the groups using knowledge and regulation of cognition scores.

GENERAL DISCUSSION

The purpose of the present research was to validate the Metacognitive Awareness Inventory (MAI). We focused on three related issues: (a) whether there was empirical support for the two-component view of metacognition, (b) whether the two components were related to each other, and (c) whether either of the components was related to empirical measures of cognitive and metacognitive performance.

Both experiments strongly supported the two-component model of metacognition (Brown, 1987; Flavell, 1987; Jacobs & Paris, 1987). The forced two-factor solutions observed in these experiments corresponded closely to knowledge and regulation of cognition. Knowledge of cognition measured an awareness of one's strengths and weaknesses, knowledge about strategies and why and when to use those strategies. Regulation of cognition measured knowledge about planning, implementing, monitoring, and evaluating strategy use. The internal consistency of these scales was excellent, ranging from .93 to .88. In addition, there was a high degree of similarity between factor loadings across experiments.

In contrast, neither experiment supported the multiple subcomponents view of metacognition. Experiment 1 reported six intercorrelated factors that did not correspond closely to any of the eight subprocesses described earlier (e.g., declarative, procedural, and conditional knowledge). In addition, most of these factors had marginally acceptable internal consistency values (i.e., $\alpha < .70$).

Regarding our second question, both experiments reported a statistically significant relationship between knowledge and regulation of cognition (i.e., $r = .54$ and .45, respectively). This finding is consistent with many previous theoretical accounts of metacognition (cf. Brown, 1987; Flavell, 1987). However, although the magnitude of these correlations indicates that knowledge and regulation of cognition are related, we found little evidence that they share a compensatory relationship. For example, in Experiment 2, the knowledge of cognition factor was related to higher test performance while the regulation of cognition factor was not. This suggests that each component makes a unique contribution to cognitive performance.

Regarding our third question, Experiment 2 reported a number of statistically significant relationships among the MAI and measures of metacognitive awareness and performance. One was the correspondence between pre-test judgments of monitoring ability and the MAI. The results of this analysis found that knowledge of cognition was related to pre-test judgments. This outcome was expected given that pre-test judgments tap knowledge of one's own memory. A second finding was that pre-test judgments were related positively to test performance. Individuals who expressed a greater degree of confidence in their monitoring ability performed better overall. A third finding was the significant relationship between the MAI and test performance. Individuals who scored high on the knowledge of cognition factor also scored high on the reading comprehension test. This finding reveals that the MAI provides valuable predictive information about subsequent performance.

In contrast, there was no statistical relationship between monitoring accuracy and the MAI or between monitoring accuracy and pre-test judgments. These findings were unexpected since knowledge about regulatory aspects of cognition was expected to improve monitoring accuracy. One reason this relationship was not observed is that a high degree of within-group variability occurred. Thus, even though we observed the expected trend in monitoring accuracy across groups, these differences were not statistically significant. A second explanation is that a test such as reading comprehension minimizes individual differences in monitoring accuracy because it relies on skills that are fully automated in the vast majority of college-age readers. In support of this claim, Corkill and Koshida (1993) found that metacognitive awareness assessed via verbal interviews was related to monitoring accuracy on a math test, but was not related to monitoring accuracy on a reading comprehension test.

Together, the findings reported here indicate that the MAI provides a reliable initial test of metacognitive awareness among older students. Using the MAI may be a helpful strategy for planning subsequent metacognitive training. The MAI also may help identify lower performing students who frequently display comprehension monitoring deficiencies that can be remediated using a variety of instructional strategies (Pressley & Ghatala, 1990).

In addition, further studies are needed to investigate the MAI's convergent and divergent validity. One important question is the relationship between the MAI and on-line verbal reports of metacognitive awareness (Corkill & Koshida, 1993; Swanson, 1990). Finding a high correlation between the two may eliminate the need to conduct time consuming, on-line interviews of metacognitive awareness. Another issue is how the MAI is related to performance and monitoring accuracy when engaged in well-defined reasoning tasks such as syllogistic and analogical reasoning or ill-defined tasks such as argumentative thinking (Kuhn, 1991). We believe the predictive validity of the MAI will increase as cognitive tasks increase in difficulty. One reason is that more individual variation occurs on difficult tasks. A second reason is that metacognitive awareness may play a greater role in the performance of complex tasks than highly automated ones like the reading comprehension test used in Experiment 2. Additional research is needed as well that compares the MAI to related inventories such as the *LASSI* (Weinstein, Palmer, & Schulte, 1987) and *MSLQ* (Pintrich, Smith, Garcia, & McKeachie, 1991). For example, the knowledge of cognition factor may be highly correlated to Pintrich *et al.*'s strategy-use factor. If so, given the MAI's high reliability, it would provide a more reliable estimate of this dimension than the MSLQ. The degree to which divergent validity exists among these instruments is an important step in determining whether the MAI measures aspects of metacognitive awareness that other instruments do not.

APPENDIX A

Items Included in the Metacognitive Awareness Inventory

1. I ask myself periodically if I am meeting my goals. (M)

2. I consider several alternatives to a problem before I answer. (M)
3. I try to use strategies that have worked in the past. (PK)
4. I pace myself while learning in order to have enough time. (P)
5. I understand my intellectual strengths and weaknesses. (DK)
6. I think about what I really need to learn before I begin a task. (P)
7. I know how well I did once I finish a test. (E)
8. I set specific goals before I begin a task. (P)
9. I slow down when I encounter important information. (IMS)
10. I know what kind of information is most important to learn. (DK)
11. I ask myself if I have considered all options when solving a problem. (M)
12. I am good at organizing information. (DK)
13. I consciously focus my attention on important information. (IMS)
14. I have a specific purpose for each strategy I use. (PK)
15. I learn best when I know something about the topic. (CK)
16. I know what the teacher expects me to learn. (DK)
17. I am good at remembering information. (DK)
18. I use different learning strategies depending on the situation. (CK)
19. I ask myself if there was an easier way to do things after I finish a task. (E)
20. I have control over how well I learn. (DK)
21. I periodically review to help me understand important relationships. (M)
22. I ask myself questions about the material before I begin. (P)
23. I think of several ways to solve a problem and choose the best one. (P)
24. I summarize what I've learned after I finish. (E)
25. I ask others for help when I don't understand something. (DS)
26. I can motivate myself to learn when I need to. (CK)
27. I am aware of what strategies I use when I study. (PK)
28. I find myself analyzing the usefulness of strategies while I study. (M)
29. I use my intellectual strengths to compensate for my weaknesses. (CK)
30. I focus on the meaning and significance of new information. (IMS)
31. I create my own examples to make information more meaningful. (IMS)
32. I am a good judge of how well I understand something. (DK)
33. I find myself using helpful learning strategies automatically. (PK)
34. I find myself pausing regularly to check my comprehension. (M)
35. I know when each strategy I use will be most effective. (CK)
36. I ask myself how well I accomplished my goals once I'm finished.

(E)

37. I draw pictures or diagrams to help me understand while learning. (IMS)
38. I ask myself if I have considered all options after I solve a problem. (E)
39. I try to translate new information into my own words. (IMS)
40. I change strategies when I fail to understand. (DS)
41. I use the organizational structure of the text to help me learn.
42. I read instructions carefully before I begin a task. (P)
43. I ask myself if what I'm reading is related to what I already know. (IMS)
44. I reevaluate my assumptions when I get confused. (DS)
45. I organize my time to best accomplish my goals. (P)
46. I learn more when I am interested in the topic. (DK)
47. I try to break studying down into smaller steps. (IMS)
48. I focus on overall meaning rather than specifics. (IMS)
49. I ask myself questions about how well I am doing while I am learning something new. (M)
50. I ask myself if I learned as much as I could have once I finish a task. (E)
51. I stop and go back over new information that is not clear. (DS)
52. I stop and reread when I get confused. (DS)

Note. DK, declarative knowledge; PK, procedural knowledge; CK, conditional knowledge; P, planning; IMS, information management strategies; M, monitoring; DS, debugging strategies; and E, evaluation.

APPENDIX B

Operational Definitions of Component Categories

Knowledge of Cognition

1. Declarative knowledge: knowledge about one's skills, intellectual resources, and abilities as a learner.
2. Procedural knowledge: knowledge about *how* to implement learning procedures (e.g., strategies).
3. Conditional knowledge: knowledge about *when* and *why* to use learning procedures.

Regulation of Cognition

1. Planning: planning, goal setting, and allocating resources *prior* to learning.
2. Information management: skills and strategy sequences used on-line

to process information more efficiently (e.g., organizing, elaborating, summarizing, selective focusing).

3. Monitoring: assessment of one's learning or strategy use.

4. Debugging: strategies used to correct comprehension and performance errors.

5. Evaluation: analysis of performance and strategy effectiveness after a learning episode.

REFERENCES

ARTZT, A. F., & ARMOUR-THOMAS, E. (1992). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction*, **9**, 137–175.

BAKER, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review*, **1**, 3–38.

BROWN, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65–116). Hillsdale, NJ: Erlbaum.

BROWN, J. E., BENNETT, J. M., & HANNA, G. (1981). *The Nelson–Denny Reading Test, Form E.* Chicago, IL: The Riverside Publishing Company.

CORKILL, A. J., & KOSHIDA, D. T. (1993). *Level of metacognitive awareness and calibration of performance: Strategic knowledge makes a difference.* Paper presented at the annual conference of the American Educational Research Association, Atlanta, GA.

FLAVELL, J. H. (1987). Speculations about the nature and development of metacognition. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 21–29). Hillsdale, NJ: Erlbaum.

GARNER, R., & ALEXANDER, P. A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologists*, **24**, 143–158.

GLENBERG, A. M., & EPSTEIN, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition*, **15**, 84–93.

JACOBS, J. E., & PARIS, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, **22**, 255–278.

KUHN, D. (1991). *The skills of argument.* New York: Cambridge University Press.

LEONSARIO, R. J., & NELSON, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 464–470.

MARKMAN, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, **48**, 986–992.

PINTRICH, P. R., SMITH, D. A. F., GARCIA, T., & MCKEACHIE, W. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ).* Ann Arbor, MI: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

PRESSLEY, M., & GHATALA, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist*, **25**, 19–33.

SWANSON, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, **82**, 306–314.

WEINSTEIN, C. E., PALMER, D. R., & SCHULTE, A. C. (1987). *LASSI: Learning and study strategies inventory.* Clearwater, FL: H & H Publishing Company.