Title: Learning and individual differences.

Article: Coutinho, S "Metacognition, need for cognition and use of explanations during ongoing learning and problem solving"

Vol: 15  No.: 4  Pages: 321-337  Date: 2005

Verified: SFX@NIUWorldCat CODEN: LIDIEI Desc: v.: Type: Serial, Internet Resource 1041-6080

Borrowing Notes:
ShipTo: Information Delivery Services/Northern Illinois University Libraries/DeKalb, IL 60115

E-delivery
Addr: (815) 753-2003  Ariel 151.156.159.48
Ship Via: Ariel, Fax, Mail, ILDS-RT1-NI

Ship Via: Ariel, Fax, Mail,

Return To:
Library
Antioch University New England
40 Avon St
Keene, NH 03431-3516

Ship To:
Information Delivery Services
Northern Illinois University Libraries
DeKalb, IL 60115

ILL: 37085274  Lender: AN@
Req Date: 11/13/2007  OCLC #: 18810223
Patron: britt, ann
Author:
Title: Learning and individual differences.
Article: Coutinho, S "Metacognition, need for cognition and use of explanations during ongoing learning and problem solving"
Vol.: 15  No.: 4  Date: 2005  Pages: 321-337
Verified: SFX@NIUWorldCat CODEN: LIDIEI Desc:
Maxcost: $25IFM  Due Date:
Lending Notes:
Bor Notes:
Metacognition, need for cognition and use of explanations during ongoing learning and problem solving

Savia Coutinho *, Katja Wiemer-Hastings, John J. Skowronski, M. Anne Britt

Northern Illinois University, United States

Received 22 March 2005; received in revised form 16 June 2005; accepted 21 June 2005

Abstract

Two studies assessed whether: (1) high levels of task-relevant metacognition would be related to good task performance; (2) some kinds of feedback (e.g., explanations) would improve task-relevant metacognition (and hence, performance) more than other kinds of feedback; and (3) some kinds of people would be more likely to seek out and use this beneficial feedback than others. Results showed that: (1) students were able to better estimate their task performance with increasing experience at the task; (2) students who provided better estimates of their task success were more successful at the task; (3) students high in need for cognition sought out problem explanations more often than students low in need for cognition; but (4) students who scored high in trait metacognition did not seek out problem explanations more often than students who scored low in trait metacognition; (5) students who were high in need for cognition performed better at the task than those who were low in need for cognition, and (6) the receipt of problem explanations was only weakly related to high levels of task performance, if at all. The implications of these results are discussed.

© 2005 Elsevier Inc. All rights reserved.

Keywords: metacognition; need for cognition; learning

Previous research has demonstrated that when people are trying to learn, feedback that includes explanations of problem solutions (and not just the solutions themselves) improves task performance. Studies suggest that explanatory-based feedback has this effect, in part, because it enables individuals to better monitor their performance (Nietfeld & Schraw, 2002) and to discover,
implement and assess effective learning strategies (Kohler, 2002) that lead to increased proficiency levels (Talbot, 1997). Thus, one potential consequence of providing an explanation for how a problem might be solved is that it affects an individual’s metacognition: individuals may become more conscious of the problem-solving process and about ways to strategize and monitor progress during that process.

However, in typical experiments, problem explanations are often imposed on people, regardless of whether they want them or not. In real-world settings, the receipt of such explanations might depend on the extent to which an individual is motivated to seek explanations out and to use them once they are received. One individual differences variable that might be related to this tendency is trait metacognition. Individuals who are high in trait metacognition are those who are thought to be especially likely to think about learning, to ponder strategies for learning, and to monitor their learning progress. A second individual differences variable that might be related to the tendency to seek out and use problem explanations is need for cognition. Individuals who are high in need for cognition are thought to be especially motivated to acquire information about, and to think about, the world around them. The research reported in this article investigates whether individuals who are high in either of these two characteristics will, indeed, be more likely to seek out explanations for problems that they encounter in the world than those who are low in these characteristics, whether such behavior is related to metacognition, and whether metacognition is related to task performance.

In the sections that follow, we will first review work on metacognition. In particular, we will briefly discuss how metacognition might be related to task performance and how metacognition might be affected by problem explanations. A second section will review the construct of need for cognition and will attempt to relate need for cognition to the seeking of performance feedback and to problem-solving performance. Subsequent sections will describe our research and its implications.

1. Metacognition

1.1. Metacognition and performance

Metacognition is an umbrella term that subsumes metacomprehension, self-monitoring, metacognitive monitoring, and self-directed learning. That is, metacognition refers to higher-order mental processes that are often involved in learning — making plans for learning, monitoring learning rates, and predicting performance (Dunlosky & Thiede, 1998). The possession of good metacognitive abilities is thought to improve performance (Dunning, Johnson, Ehrlinger, & Kruger, 2003), and data suggest that many individuals lack such abilities. For example, a meta-analysis by Mabe and West (1982) revealed that people often misperceive their own ability levels. Kruger and Dunning (1999) showed that in academic settings, such miscalibration is especially characteristic of poor academic performers. The implication is that good metacognition (in the form of good calibration) leads to good academic performance.

A similar conclusion comes from Everson and Tobias (1998), who found that metacognitive ability (as measured by metacognitive word knowledge, or KMA) was a predictor of word test performance. Their results suggested that when trying to learn word skills, students high in KMA were especially likely to realize what they did not know and to take steps to remedy the problem (e.g., engage in relearning strategies). The ability to accurately engage in metacognitive monitoring and to engage in
subsequent self-regulation (Carver & Scheier, 1990) has been similarly implicated as a key to academic and problem-solving success by a number of other researchers (Dunlosky & Thiede, 1998; Thiede, Anderson, & Therriault, 2003).

1.2. Metacognitive training and feedback

Another line of metacognition research has explored the possibility that providing feedback about performance can create long-term improvements in metacognition, which can then positively influence performance. For example, Thiede et al. (2003) asked participants to read text passages, to rate their comprehension of each passage, and to answer questions about each. Participants received overall feedback regarding test performance. The results of the study showed that feedback positively influenced subsequent test performance.

However, some kinds of feedback might better prompt metacognition than others. For example, in a study by Kruger and Dunning (1999), participants solved problems based on the Wason selection task and estimated their performance. As expected, poor performers overestimated their logic reasoning ability. However, after the test some participants were given training in logic and reasoning and were given their own tests back. They indicated how many logic problems they had solved correctly and incorrectly and re-estimated their ranking relative to their peers. The combination of training and feedback significantly improved students’ metacognitive ability — they showed an improvement in their calibration judgments. Moreover, such enhancement has been linked to improved task performance. For example, an experiment conducted by Leasure (1997) showed that a combination of metacognitive training and feedback was found to improve performance predictions in a remedial math course, as well as improving actual performance. Similar results have been obtained by Kohler (2002), Nietfeld and Schraw (2002), and Talbot (1997).

The implications of these findings are clear. Good task performance may result from good metacognition; thus, individuals with good metacognitive skills may perform better than those who do not. One way to improve metacognitive skills for a particular task may be to provide individuals explanations of representative problem solutions.

2. The need for cognition

Cohen, Stotland, and Wolfe (1955) conceptualized the need for cognition as “a need to structure relevant situations in meaningful, integrated ways. It is the need to understand and make reasonable the experimental world” (p. 291). Cacioppo and Petty (1982) revised this idea, envisioning the need for cognition as “the tendency to engage in and enjoy effortful cognitive activity” (p. 116). They conceived of need for cognition as a tendency and not an invariant trait or disposition. Hence, the need for cognition might be affected by situational circumstances, such as time pressure or decision importance.

However, Cacioppo and Petty also surmised that there were durable individual differences in need for cognition. Individuals who are high in the need for cognition are thought to be intrinsically motivated to think and to enjoy complex cognitive tasks. In contrast, individuals who are low in need for cognition are described as “cognitive misers” (Taylor, 1981) who choose not to expend energy on cognitive activities. (p. 127).
To measure this individual difference, Cacioppo, Petty, and Kao (1984) developed an 18-item Need for Cognition Scale (NCS) (a revision of an earlier 34-item scale) that has desirable psychometric properties (also see Forsterlee & Ho, 1999; Sadowski, 1993). The need for cognition construct and the individual-differences measure of need for cognition appear to have considerable utility and validity. There is now a vast body of literature that shows how need for cognition affects responding in areas ranging from social and cognitive psychology to medicine, journalism and law (Cacioppo, Petty, Feinstein, & Jarvis, 1996).

2.1. Need for cognition and academic performance

Of particular interest to the present article are studies illustrating that the need for cognition is related to academic performance. For example, Leone and Dalton (1988) administered the 18-item NCS to 87 students in an introductory social psychology class. They found that students high in need for cognition received higher grades on material that required high cognitive effort than students low in need for cognition. In a related study, Sadowski and Gulgoz (1996) examined how the need for cognition affected elaborative processing in academic performance. One characteristic of elaborative processing is the ability to generate explanations relevant to target concepts. Undergraduate students in the study completed the NCS and selected advertisements, cartoons or other materials that reflected important concepts covered in their course. They then described the processes they used in selecting the particular material and explained how the material they selected was relevant. Elaborative processing was assessed by examining the explanations students provided for complexity and the presence of relational networks. Students high in need for cognition generated more complex and elaborate explanations of target concepts and had better grades on tests relating to their coursework than students low in need for cognition.

Clearly, then, individuals with high need for cognition perform better in intellectually challenging tasks. However, the need for cognition may play itself out in different ways then metacognition. Good metacognition is likely applied by individuals directly to improve the outcome of a problem solving process, and may thus to some extent depend on task interest and motivation to perform well. Need for cognition may influence task performance independently because it engages individuals in the problem solving process. As such, individuals with high need for cognition are also likely to seek out and to benefit from problem solutions and explanations, because they may be particularly interested in why a problem solution works and how such rationales can be applied to their future problem-solving activities. While they may not be especially likely to engage in the same self-regulatory activity that characterizes those who are high in metacognition levels, their general interest in thinking may similarly lead to the use of learning strategies that can effectively enhance performance.

3. Overview of experiments

The experiments for this study were designed to weave together the themes described in the sections above. Students in these experiments worked to solve a set of analytical problems from the Graduate Record Exam (GRE). An indicator of metacognitive performance (accuracy calibration) was obtained after each problem solution was submitted. After establishing a baseline performance in a first block of items, students received two forms of feedback on a second block of items. In Experiment 1, at the end of each problem, students could choose to either view the solution with an explanation, or to view just the
solution. In Experiment 2, all students were exposed to solutions with explanation. In both experiments, performance on the problems in the second block was examined to test whether students’ need for cognition, level of trait metacognition, or the tendency to obtain problem explanations affected task performance.

In Experiment 1, we expected that those who scored high in need for cognition, or who indicated a strong tendency to engage in metacognitive activity, would be more likely to request answers that also explained the solutions to problems. We also expected that the provision of such problem solutions would enhance calibration accuracy. In Experiment 1 we expected enhanced calibration accuracy primarily for those individuals who were high in need for cognition or who were high in trait metacognition. However, in Experiment 1 such effects could occur because of selective exposure to the problem explanations. Experiment 2 provided explanations for all problems, thus, resultant individual differences in calibration accuracy could not be explained by the selective tendency to seek out explanations. Instead, such differences may reflect interactive effects of individual differences and explanations on metacognitive activity. Finally, we expected enhanced performance to follow from enhanced metacognition. That is, we expected that individuals who received problem explanations should be particularly likely to evince enhanced performance. Such enhanced performance should be especially evident in those who are particularly likely to seek out and use problem explanations: individuals who are high in trait metacognition or high in need for cognition.

4. Experiment 1

4.1. Method

4.1.1. Participants

Participants were 250 undergraduate students enrolled in an introductory psychology course at Northern Illinois University. Ages ranged from 16 years to 46 years ($M = 19.1$, $SD = 2.47$). Students participated in this experiment to fulfill class requirements. Data from thirteen participants were dropped from the analyses due to abnormally rapid responses to the problem-solving task. Our judgment was that such rapid response times were indicative of subjects who were not performing the task as instructed. Thus, only data from 237 participants (115 women and 122 men) who appeared to be engaging in the task in good faith were included in the analyses.

4.1.2. Individual-difference measures

In the course of completing the study, participants in Experiment 1 completed two inventories: the 18-item NCS (Cacioppo et al., 1984) and the TMC (Hong & O’Neil, 2001). Participants responded to each of the 18 NCS items on a 9-point Likert-type scale anchored by strongly disagree on the low end (1), strongly agree at the high end (9), and neither agree nor disagree at the midpoint (5). As directed by Cacioppo et al., after reverse scoring some items, a total need for cognition score for each student was computed by summing responses to all items. Students responded to each item on the TMC using a four-point response scale. The response options were: (1) almost never, (2) sometimes, (3) often, and (4) almost always. A total TMC score for each student was computed by adding the scores for all items.
4.1.3. Problems to be solved

Students also attempted to complete up to 20 GRE-Analytical items. These items were taken from the test preparation guide entitled *GRE: Practicing to take the General Test* (Educational Testing Services, 1990). This book provides an explanation for each item’s answer, and in Experiment 1, these explanations were used verbatim as the problem explanations for the experiment. In addition, the book also provides an estimate of the difficulty of each item. This estimate is based on the percentage of individuals who typically answer an item correctly. These data were used to select items for the experiment. Each set of 10 GRE items given to participants was composed of 2 moderately difficult items, 4 difficult items and 4 extremely difficult items (for a sample item and problem explanation, see Appendix A).

4.1.4. Procedure

All instructions and materials were presented via personal computer. To establish a performance baseline, participants first completed one set of 10 GRE analytical items. Participants read each GRE analytical item and responded by choosing the correct answer from a set of five possible answers. Responses to each item were timed to ensure that participants faithfully attempted to solve the problems. Immediately after making an answer to each given item, participants reported their perception of the accuracy of their answer by indicating on a five-point scale whether they thought that their answer was (1) certainly incorrect, (2) probably incorrect, (3) uncertain, (4) probably correct or (5) certainly correct.

Immediately following this first set of items, participants attempted a second set of 10 GRE analytical items. As before, subjects read the problem, chose an answer, and estimated their accuracy. However, in this set, after responding to this item, participants then had the option of: (a) obtaining the problem’s answer accompanied by a short explanation of problem’s solution, or (b) simply obtaining the correct answer to the problem.

Finally, after completing the second set of 10 GRE items, participants completed the NCS and TMC. They were then debriefed and dismissed.

4.1.5. Performance and calibration measures

4.1.5.1. Performance accuracy: Participants’ performance on each set of items was assessed by adding the number of items correctly answered in each set. Hence, values on this measure could range from 0 to 10 for each set.

4.1.5.2. The accuracy prediction confidence score. The Accuracy Prediction Confidence Score (APC) was derived for each participants’ from their responses to each GRE item. If an individual’s judgment of an item was correct, the APC score assigned to that item was simply the participant’s confidence rating for that item. On the other hand, if the response was incorrect, the individual’s judgment was reverse-scored. For example, if an individual responded correctly to an item and rated his or her confidence in the response as 5 (certainly correct), an APC score of 5 would be assigned to that individual for that for that item — an instance of good calibration. However, if an individual responded incorrectly to a GRE item and rated his or her confidence as 5, this individual would obtain an APC score of 1 for the item — an instance of poor calibration.

An overall APC score was computed for participants on each set of GRE items by summing up the ten APC scores on that set of GRE items. Therefore, an individual with excellent metacognition who
responded to all GRE items correctly and rated his or her accuracy prediction accuracy as certainly correct (5) could potentially receive an APC score of 50 for that set of GRE items (5 points for each of 10 items). An individual with poor metacognition who responded to every GRE item incorrectly but rated his or her accuracy for each item as certainly correct (5) would receive a total score of 10 (1 point for each of 10 items).

4.2. Results

4.2.1. Descriptive

Table 1 presents descriptive statistics for the NCS and TMC scores, accuracy scores for each set of GRE items, and APC scores for each set of items. The means for the NCS and TMC scores reflect values that are expected given prior research with these instruments. The standard deviations reflect reasonable variation among individuals on these scales.

However, the data for performance accuracy are less encouraging. The means are quite low and there is relatively little variability in the data. Such results suggest the possibility of a floor effect in performance: we may have chosen problems that were simply too difficult for our population. At the same time, peoples’ calibration scores fell in mid-scale and there was reasonable variation in calibration scores. This suggests that although everyone did poorly on the test, some people were more attuned to their poor performance than others.

4.2.2. Frequency of viewing problem explanations versus viewing only correct answers

It was expected that participants who were high in need for cognition or in trait metacognition would be especially likely to request explanations for the GRE problem answers. To test this prediction, a median split on NCS scores was used to assign participants to one of two need for cognition groups (low need for cognition participants and high need for cognition participants). Scores on the TMC measure were similarly used to classify individuals as high in trait metacognition or as low in metacognition. This created four groups: people high in both need for cognition and in trait metacognition (N=66); people who were high in need for cognition and low in trait metacognition (N=59); people who were low in need for cognition and high in trait metacognition (N=45); and people who low in both need for cognition and trait metacognition (N=67).

Consistent with the predictions, results of the analysis show that individuals who were high in need for cognition requested problem explanations more often (M=2.18) than those who were low in need for cognition (M=1.41), F (1,233)=4.59, p<.05. However, neither the trait metacognition main effect nor

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCS</td>
<td>104.48</td>
<td>18.53</td>
<td>47.00</td>
<td>158.00</td>
</tr>
<tr>
<td>TMC</td>
<td>91.65</td>
<td>13.93</td>
<td>60.00</td>
<td>128.00</td>
</tr>
<tr>
<td>GRE 1</td>
<td>2.55</td>
<td>1.63</td>
<td>0.00</td>
<td>8.00</td>
</tr>
<tr>
<td>GRE 2</td>
<td>2.46</td>
<td>1.47</td>
<td>0.00</td>
<td>8.00</td>
</tr>
<tr>
<td>APC for GRE 1</td>
<td>27.87</td>
<td>4.81</td>
<td>10.00</td>
<td>43.00</td>
</tr>
<tr>
<td>APC for GRE 2</td>
<td>28.80</td>
<td>4.54</td>
<td>11.00</td>
<td>43.00</td>
</tr>
</tbody>
</table>
the Need for Cognition x Trait Metacognition interaction were significant ($F (1, 233) = 2.06, p > .16$ and $F (1, 233) = .187, p > .67$, respectively).\footnote{All analyses were also conducted using regression procedures, which allow predictor variables to be used in their fully continuous form. We have opted to report the results of the ANOVA analyses because they allow for simpler data presentation. However, the regression analyses were valuable in showing that, in accord with what would be expected from the literature, effects that merely approached significance using the median split methods were statistically reliable when the full range of the continuous predictors were used. Hence, we are comfortable reporting median split results that merely approach significance because those effects are significant using the regression methods.}

4.2.3. Calibration (APC) scores

A second area of interest concerned the extent to which participants were able to correctly gauge their own performance as reflected in the APC scores. Two separate analyses explored these scores.

4.2.3.1. APC scores as a function of need for cognition, trait metacognition and task experience. In one analysis the APC score tallied for each participant for each set of GRE items was entered into a Need for Cognition $\times$ Trait Metacognition $\times$ Trial Block (first versus second) mixed ANOVA. It was expected that calibration performance would be related to task experience — as the task progressed one should be better able to judge whether one was successful or not. One question of interest was whether this expected increase would be affected by a participant’s need for cognition or trait metacognition.

The results of this analysis showed that, as expected, participants’ calibration scores significantly increased across trial blocks (block 1 $M = 27.87$, block 2 $M = 28.80$), $F (1, 233) = 7.25, p < .003$. While this increase was greater for participants with high trait metacognition (block 1 $M = 27.44$; block 2 $M = 28.60$) than for those with low trait metacognition (block 1 $M = 28.29$; block 2 $M = 29.00$), this interaction was not significant $F (1, 233) = 1.50, p > .23$. This lack of statistical significance characterized all other effects in the analysis.

4.2.3.2. APC scores as a function of viewing problem solutions. A second analysis explored the idea that viewing problem solutions improved task-related metacognition as reflected in participants’ APC scores. To explore this idea, a regression analysis was conducted in which each participant’s APC score from the second set of GRE items was predicted by the number of problem solutions that an individual had requested. Although viewing problem solutions was positively related to the APC scores ($b = .152$), the relation between the two variables was not significant, $F (1, 235) = 2.02, p > .16$.

4.2.4. GRE item performance

An additional question of interest in this study was the extent to which task experience, need for cognition, trait metacognition, explanation-viewing, or calibration might be related to problem solving performance. Three analyses explored this issue.

4.2.4.1. Performance scores as a function of need for cognition, trait metacognition and task experience. One analysis examined the GRE performance scores as they were related to task experience, need for cognition and trait metacognition. The total number of GRE items correctly solved in each item set was entered into a mixed ANOVA with need for cognition and trait metacognition as the between-participant variables and trial block (first vs. second) as the within-participant variable.
It was expected that task experience would be related to improvements in task performance, and that high need for cognition or high trait metacognition would be related to high levels of performance or greater improvement across trial blocks. The results showed that, averaging across trial blocks, high need for cognition participants ($M=2.68$) outperformed low need for cognition participants ($M=2.35$), $F (1, 233)=4.46$, $p<.04$. However, the results also showed that participants who were high in need for cognition performed worse across trial blocks (block 1 $M=2.93$; block 2 $M=2.43$) while those who were low in need for cognition did not (block 1 $M=2.19$; block 2 $M=2.50$), $F (1, 233)=10.45$, $p<.001$. No other significant effects emerged from the analysis.

4.2.4.2. Performance scores as a function of APC score. A second analysis examined whether participants’ calibration scores predicted task performance. To conduct these analyses a measure of calibration (APC score summed across trial blocks) was used to predict performance (total correct GRE items across both trial blocks). Results showed that subjects who were better calibrated did, indeed, perform better on the GRE items ($b=.099$, $F (1, 235)=25.79$, $p<.001$).

4.2.4.3. Performance scores as a function of explanation-viewing. A third analysis examined the extent to which task performance feedback, task-relevant metacognition, and problem-solving performance. The experiment tested the predictions that (1) high levels of task-relevant metacognition would be related to good task performance; (2) some kinds of feedback (e.g., explanations) would improve task-relevant metacognition (and hence, performance) more than other kinds of feedback; and that (3) individual differences would account for the variability in the tendency to seek out and use feedback.

The data that were obtained from Experiment 1 were only partially supportive of these ideas. These results showed that a measure of one task-relevant metacognitive ability, calibration (as measured by APC), increased with feedback-inclusive task experience. In addition, calibration performance was positively related to task performance. These results are consistent with the idea that task performance is improved when task-relevant metacognition is improved.

5. Discussion

Experiment 1 explored links between task performance feedback, task-relevant metacognition, and problem-solving performance. The experiment tested the predictions that (1) high levels of task-relevant metacognition would be related to good task performance; (2) some kinds of feedback (e.g., explanations) would improve task-relevant metacognition (and hence, performance) more than other kinds of feedback; and that (3) individual differences would account for the variability in the tendency to seek out and use feedback.

The data that were obtained from Experiment 1 were only partially supportive of these ideas. These results showed that a measure of one task-relevant metacognitive ability, calibration (as measured by APC), increased with feedback-inclusive task experience. In addition, calibration performance was positively related to task performance. These results are consistent with the idea that task performance is improved when task-relevant metacognition is improved.

Similar implications come from the fact that a personal characteristic that should be related to better task-relevant metacognition, need for cognition, was also positively related to better task performance. Consistent with ideas about feedback and metacognition, high need for cognition individuals were more likely than low need for cognition individuals to acquire problem explanations that are thought to improve metacognitive abilities. However, other results question whether the better performance of high need for cognition individuals occurred because of their better calibration: The results of our analyses also showed that high need for cognition was not related to the possession of a high APC score. In addition, an individual-differences measure of trait metacognition, the TMC score, was not related to calibration, nor to information-seeking, nor was it related to task performance.
One might expect that the opportunity to acquire problem descriptions that could improve task-relevant metacognition would have especially increased the performance of high need for cognition participants across trial blocks. Yet, despite the fact that they were more likely to seek problem explanations, the performance of high need for cognition students actually went down from trial block 1 to trial block 2.

It is possible that this decrease was caused by the greater number of problem explanations acquired by those who were high in need for cognition. However, those explanations should have improved both task-relevant knowledge and task-relevant metacognition, so it struck us as unlikely that they should have decreased task performance. Instead, we wondered whether this performance decrease across trial blocks was somehow an unintended consequence of giving individuals a choice about the kinds of feedback that they could receive. That is, we wondered whether high need for cognition individuals might be expending cognitive resources in their choice-making that would otherwise be steered toward problem-solving. The low thinking-motivation associated with low need for cognition individuals may render them immune to the side-effects of such choice-making.

We decided to clarify this puzzling finding in Experiment 2 by eliminating the choice-making option that was presented in Experiment 1. In Experiment 2, all students were given explanations for all problem solutions in trial block 2. Hence, if choice-making was selectively interfering with the problem-solving activity of high need for cognition participants, one would expect this interference to be absent in Experiment 2. In addition to addressing this question, Experiment 2 provides an opportunity to re-examine some of the main questions that drive this article and to replicate some of the outcomes that emerged in Experiment 1.

6. Experiment 2

6.1. Method

6.1.1. Participants
Participants were 187 undergraduates enrolled in an introductory psychology course at Northern Illinois University. Ages ranged from 18 years to 35 years ($M=19.15$, $SD=1.92$). Students participated in this experiment to fulfill class requirements. Data from seven participants were dropped from the analyses due to their abnormally rapid responses to the problem-solving task. As in Experiment 1, such hasty responses were believed to be indicative of participants who were not performing the task as instructed. Therefore, only data from 180 participants (94 women and 86 men) who appeared to be performing the task as instructed were included in the analyses.

6.1.2. Measures
Scales, GRE items, and problem explanations used in Experiment 2 were the same as those used in Experiment 1.

6.1.3. Procedure
Experiment 2 was similar to Experiment 1 — there was only one difference. In Experiment 2, on the second set of 10 GRE analytical items, all participants read the problem, chose an answer, estimated their accuracy and were then exposed to the problem’s answer accompanied by an explanation of the
problem’s solution. Participants did not have the option of obtaining either the correct answer or the correct answer and an explanation, as they did in Experiment 1.

6.2. Results

6.2.1. Descriptives

Table 2 presents descriptive statistics for the NCS and TMC scores, accuracy scores for each set of GRE items, and APC scores for each set of GRE items. The pattern of means and variability mirror the results from Experiment 1, with expected means and good variation for the NCS and TMC, low overall performance and variability in performance on GRE items, as well as acceptable calibration scores. Again, there was sufficient variation in calibration scores to indicate that some students were more aware of their poor performance than other students.

6.2.2. Calibration (APC) scores

One area of interest is the extent to which participants were able to correctly estimate their task performance. This capability is reflected in participants’ APC scores. In Experiment 2 we were interested in whether these APC scores would be influenced by participants’ experience with the task, their metacognition level, or their need for cognition. As in Experiment 1, in preparation for this analysis we conducted median splits on the scores derived from the NCS and TMC. This created four groups: people high in both need for cognition and in trait metacognition (N=58); people who were high in need for cognition and low in trait metacognition (N=33); people who were low in need for cognition and high in trait metacognition (N=30); and people who low in both need for cognition and trait metacognition (N=59).

The APC scores for each participant for each trial block was entered into a Need for Cognition × Trait Metacognition × Trial Block (first versus second) mixed ANOVA. The results of this analysis showed that, as in Experiment 1, people became better calibrated across trial blocks (block 1 M=26.88; block 2 M=28.53), F (1,176)=15.39, p<.001. Hence, experience with the task did appear to improve the extent to which people were able to estimate their own performance. As in Experiment 1, no other significant effects emerged from the analysis. Hence, this improvement in calibration was not moderated by either the participant’s TMC score or by the participant’s NCS score.

6.2.3. GRE item performance

We were interested in the extent to which need for cognition, trait metacognition, task experience, or calibration might be related to successful problem-solving. Two analyses were conducted to explore such relations.

Table 2
Means, standard deviations, minimum and maximum scores from Experiment 2 measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCS</td>
<td>103.18</td>
<td>21.99</td>
<td>38.00</td>
<td>160.00</td>
</tr>
<tr>
<td>TMC</td>
<td>91.48</td>
<td>14.13</td>
<td>58.00</td>
<td>134.00</td>
</tr>
<tr>
<td>GRE 1</td>
<td>2.41</td>
<td>1.66</td>
<td>0.00</td>
<td>8.00</td>
</tr>
<tr>
<td>GRE 2</td>
<td>2.72</td>
<td>1.73</td>
<td>0.00</td>
<td>8.00</td>
</tr>
<tr>
<td>APC for GRE 1</td>
<td>26.88</td>
<td>4.41</td>
<td>13.00</td>
<td>37.00</td>
</tr>
<tr>
<td>APC for GRE 2</td>
<td>28.49</td>
<td>4.59</td>
<td>14.00</td>
<td>41.00</td>
</tr>
</tbody>
</table>
6.2.3.1. **Performance scores as a function of task experience, need for cognition and trait metacognition.** One analysis investigated GRE performance scores as they were related to task experience, need for cognition and trait metacognition. The total number of GRE items solved accurately in each set of GRE items was entered into a mixed ANOVA with need for cognition and trait metacognition as the between-participant variables and trial block (first versus second) as the within-participant variable.

There was a tendency for performance to improve across trial blocks (block 1 \(M = 2.41\); block 2 \(M = 2.72\), \(F(1,176) = 3.49, p < .07\). Given that this effect was not obtained in Experiment 1, one might conclude viewing of problem explanations was a critical factor for performance increases. In Experiment 1, only roughly two out of ten explanations were viewed, whereas in Experiment 2, all participants viewed ten out of ten explanations. However, given that this effect only approaches significance, and that there was no relationship between explanation-viewing and problem-solving success in Experiment 1, it would be wise to treat this effect cautiously until replicated in future research.

Replicating results from Experiment 1, averaging across the trial blocks participants who were high in need for cognition performed better \((M = 2.86)\) than participants who were low in need for cognition \((M = 2.27)\), \(F(1,176) = 8.38, p < .003\). The data also showed that there was a tendency for students low in need for cognition to improve their performance substantially across trial blocks (block 1 \(M = 1.99\); block 2 \(M = 2.54\) while students who were high in need for cognition did not (block 1 \(M = 2.82\); block 2 \(M = 2.89\)). The interaction was not statistically reliable, \(F(1,176) = 2.01, p > .16\). This interaction was reliable in Experiment 1, but the pattern of means differed greatly — the performance of high need for cognition individuals decreased substantially across trial blocks in that experiment. The fact that this effect was eliminated in Experiment 2 suggests that our idea that choice-making activity may have been especially likely to impair the problem solving performance of high need for cognition participants in Experiment 1 was a valid one.

No other effects in this analysis were statistically reliable.

6.2.3.2. **Performance scores as a function of APC score.** A second analysis of performance scores assessed whether participants’ calibration scores predicted task performance. A measure of calibration (APC score summed across trial blocks) was used to predict performance (total GRE items across both trial blocks). As in Experiment 1, results showed that participants who were better-calibrated performed better on GRE items than participants who were less well-calibrated \((b = .131, F(1,179) = 25.39, p < .001)\).

7. **Discussion**

The data that were obtained from Experiment 2 were only partially supportive of the ideas that: (1) that high levels of task-relevant metacognition would be related to good task performance; (2) that some kinds of feedback (e.g., explanations) would improve task-relevant metacognition (and hence, performance) more than other kinds of feedback; and (3) that some kinds of people would be more likely to use this beneficial feedback than others.

The results showed that a measure of one task-relevant metacognitive ability, calibration (as measured by APC), increased with feedback-inclusive task experience. In addition, calibration performance was positively related to task performance. These results are consistent with the idea that task performance is improved when task-relevant metacognition is improved.
Similar implications come from the fact that a personal characteristic that should be related to better task-relevant metacognition, need for cognition, was also positively related to better task performance. However, other results question whether the better performance of high need for cognition individuals occurred because of their better calibration: The results of our analyses also showed that high need for cognition was related to higher task performance than that exhibited by low need for cognition individuals, even in the absence of explanatory feedback (trial block 1). Moreover, in Experiment 2 it was the low need for cognition individuals who were especially likely to benefit from the explanatory feedback. In addition, an individual-differences measure of trait metacognition, the TMC score, was not related to calibration, nor was it related to task performance.

However, Experiment 2 did appear to solve the puzzle of the decrease in performance across trial blocks exhibited by high need for cognition participants in Experiment 1. That decrease was likely due to cognitive load or interference caused by the opportunity to make choices about feedback. When that opportunity to make choices was removed in Experiment 2, the decrease in performance across trial blocks that was observed in Experiment 1 dissipated.

8. General discussion

In this paper, we set out to explore how metacognition, need for cognition, and information about problem solutions and explanations affect individuals task performance and metacognition. Task performance was measured using analytical items from the GRE. Both experiments showed that the GRE items were very difficult for most students, which greatly reduced variability in performance. As a consequence, some of our findings may not generalize to task environments that are easier to solve. At the same time, the difficulty level was an important choice because it allowed for performance improvement. Also, problem explanations were unlikely to be of much interest to participants if GRE problems were easy, because participants would not need to view the problem explanations in order to solve the problems.

8.1. Factors in task performance

Consistent with past findings, both experiments showed that individuals with high need for cognition solved more problems correctly than individuals with low need for cognition. Likewise, both experiments showed advantages for task performance for individuals who were better calibrated (as measured by the task-relevant APC scores). However, the trait metacognition measure (TMC) was not related to task performance. The discrepancy between these two forms of metacognitive awareness suggests that metacognitive skills in one area do not necessarily map onto another area. This idea is consistent with the proposal that it may be adequate to think of metacognition as a skill that is at least partially domain-dependent (Kelemen, Frost, & Weaver, 2000; Schraw, Dunkle, Bendixen, & Roedel, 1995).

Participants’ performance increased slightly from the first to the second block of trials, but only in Experiment 2. In Experiment 2, participants were consistently provided with explanations to the correct solutions, which were viewed only rarely in the first Experiment. Thus, standardizing feedback to provide explanations may be important in order to improve performance. Consistent feedback with explanations may be preferable over several feedback options for another reason: Individuals with high need for cognition performed worse in the second trial block of Experiment 1, but not Experiment 2. It is
possible that when individuals can choose the form of feedback, this choice detracts from the cognitive resources of individuals who are likely to engage in deep processing, such as those high in need for cognition. Further, the solutions to problems improved the performance of individuals with low need for cognition, who are unlikely to seek out such additional information on how to solve problems on their own. Overall then, all individuals may be more likely to benefit from consistent explanations, regardless of need for cognition levels.

8.2. Factors in explanation viewing

Experiment 1 was designed to assess whether individuals with high metacognition and high need for cognition would be more likely to seek out task-relevant information. The prediction was that such individuals would be more likely to choose solutions with explanations. Individuals with high metacognition may be motivated to seek out problem explanations in order to improve their performance; individuals with high need for cognition may be motivated to better understand the problems. These predictions were partially supported: Individuals with high need for cognition were indeed more likely to view explanations than individuals with low need for cognition. However, the results of Experiment 1 show that this does not necessarily lead to improved performance for these individuals. In fact, as mentioned above, there was a drop in performance for individuals with high need for cognition; perhaps due to the choice distracting these individuals from the main task, or due to additional demands placed on their cognitive resources.

8.3. Factors in metacognitive skill

Metacognitive skill has been shown to lead to higher performance for a variety of skills. Consistent with this view, our two experiments show that individuals who were calibrated better, performed better. Thus, to improve students’ performance on tasks, one central question is how metacognitive skill can be improved. We found that individuals became calibrated better from the first to the second trial block, which suggests that practice at doing a specific task by itself can lead to some improvement in calibration. The improvement was observed for both experiments, which suggests that it may be independent of the form of feedback. However, providing the correct solution may be critical for individuals to develop a better sense of their own performance.

9. Uncited references

Borillo, 1996
Breland, Jones & Jenkins, 1994
Brown, 1980
Cacioppo & Petty, 1984
Chi, Glaser & Rees, 1982
Cohen, 1957
Ehrlinger & Dunning, 2003
Falchikov & Boud, 1989
Garner & Alexander, 1989
Appendix A. Sample GRE analytical item and problem explanation

I. In the 1980 United States consensus, marital status was described under one of five categories: single, now married (but not separated), separated, divorced, widowed. In the category “separated”, including both those who were legally separated and those who were estranged and living apart from their spouses, one million more women than men were counted. Which of the following, if true, provide(s) or contribute(s) to an explanation for this result? (Percentage of students responding correctly: 21%).

I. There are more women of marriageable age than men of marriageable age in the United States.

II. More of the separated men than separated women in the United States could not be found by the census takers during the census.

III. Many more separated men than separated women left the United States for residence in another country.

(A) I only
(B) II only
(C) III only
(D) I and II only
(E) II and III only.

Correct Answer: E

Explanation: Proposition I is completely irrelevant to the particular imbalance. There ought to be as many men as women in this category. Proposition II on the other hand does help explain the imbalance, as does proposition III. In both cases a reason is given for why census takers did not count as many separated men as they did separated women. Therefore, (E), (II and III) is the correct answer.

References


