Athens Institute for Education and Research
ATINER

ATINER's Conference Paper Series
EDU2018-2482

Metacognitive Awareness (or Lack Thereof) During Problem Solving

Brian D. Beitzel
Professor of Educational Psychology
SUNY Oneonta
USA

An Introduction to ATINER's Conference Paper Series

Conference papers are research/policy papers written and presented by academics at one of ATINER's academic events. ATINER's association started to publish this conference paper series in 2012. All published conference papers go through an initial peer review aiming at disseminating and improving the ideas expressed in each work. Authors welcome comments

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research

This paper should be cited as follows:

Beitzel, B. D., (2018). **"Metacognitive Awareness (or Lack Thereof) During Problem Solving",** Athens: ATINER'S Conference Paper Series, No: **EDU2018-2482.**

Athens Institute for Education and Research
8 Valaoritou Street, Kolonaki, 10671 Athens, Greece
Tel: + 30 210 3634210 Fax: + 30 210 3634209 Email: info@atiner.gr URL: www.atiner.gr
URL Conference Papers Series: www.atiner.gr/papers.htm
Printed in Athens, Greece by the Athens Institute for Education and Research.
ISSN: 2241-2891
03/08/2018

Metacognitive Awareness (or Lack Thereof) During Problem Solving

Brian D. Beitzel
Professor of Educational Psychology
SUNY Oneonta
USA

Abstract

Human learners are often notably imprecise in estimating their own skills in a variety of performance domains. The present study examines undergraduate students' judgments of the mental demand they have experienced during a problem-solving task in the domain of mathematics. After receiving instruction on how to solve total-probability problems and being presented with a series of worked examples, participants were asked to solve a series of the same type of problems unassisted. Throughout the experiment, eye-tracking equipment was used to calculate estimates of mental workload using a United States-patented algorithm called the Index of Cognitive Activity (ICA). The ICA metric has been validated in a variety of performance domains and has been shown to be a reliable indicator of mental effort. After each posttest problem, participants were (a) presented with the ICA estimate for the mental workload they had just experienced while solving that problem and (b) asked whether they agreed with the estimate. For each problem, slightly more than half of the participants indicated that the ICA estimate was "about right"; the other half split between the estimate being too high or too low. Arguably, the ICA estimates are a more objective measure of the mental demand that participants experienced than their own subjective self-reflections. These findings corroborate the work of other researchers demonstrating that adult learners are not always proficient in evaluating their own performance. In addition to showing discrepancies between actual performance and self-judgments of it, this study provides evidence that the metacognitive aspects of learning are also subject to errors in self-estimation.

Keywords: metacognition, self-awareness, self-judgments

Introduction

Human learners are remarkably unskilled as judges of their own competence. This fact has been incontrovertibly established in both academic and non-academic domains (Beaudoin and Desrichard, 2011, Dunning, Heath, and Sula, 2004, Dunning, Johnson, Ehrlinger, and Kruger, 2003, Ehrlinger, Johnson, Banner, Dunning, and Kruger, 2008, Kruger and Dunning, 1999, Vazire and Mehl, 2008, Zell and Krizan, 2014). For example, Chemers, Hu and Garcia (2001) found that the correlation between first-year college students' academic self-ratings and their instructors' assessments was only 0.35. With the assumption that the instructors' ratings were reasonably accurate, this weak relationship between student and instructor ratings demonstrates a striking lack of self-awareness on the part of students regarding their academic capabilities after their first term at the university. Freund and Kasten (2012) provide further evidence documenting humans' inaccurate self-assessments. Across 154 samples in their meta-analysis, the average correlation between self-ratings of cognitive ability and formal psychological assessments of cognitive ability was 0.33 – a modest correlation at best. And Zell and Krizan (2014), in their meta-analysis across multiple domains, found an average correlation of only 0.29 between self-evaluations and actual performance. These correlations demonstrate that humans are significantly unskilled in evaluating their own capabilities.

Perhaps not surprisingly, the individuals who exhibit the most unsuccessful performances on a given task are also the least capable of evaluating their own efforts, reliably overestimating their abilities by a rather wide margin (Dunning et al., 2004, Dunning et al., 2003, Ehrlinger et al., 2008, Ferraro, 2010, Kruger and Dunning, 1999, Pennycook, Ross, Koehler, and Fugelsang, 2017). Social psychologists refer to this flattery of oneself as *self-enhancement* (Dufner, Gebauer, Sedikides, and Denissen, 2018, Hepper, Gramzow, and Sedikides, 2010, Kim, Chiu, and Zou, 2010). A multitude of studies have documented the fact that humans tend to have an artificially favorable view of themselves (i.e., a perspective of self-enhancement) across many domains (Dufner et al., 2018).

Lew, Alwis and Schmidt (2010) showed that self-assessments of low-performers do not improve over time, even in a context that provides repeated feedback from a tutor on the quality of their work. However, increasing the capabilities of unskilled individuals has been shown to concomitantly improve the accuracy of their self-assessments, at least in knowledge-based domains (Ehrlinger et al., 2008). This suggests that low-skilled individuals are not inherently incapable of assessing themselves; rather, when they lack a suitable level of expertise they do not have the knowledge base upon which to base a judgment. Thus, the least skilled individuals suffer from a dual burden: they are not only (a) the lowest-ranked performers in terms of achievement, but also (b) they lack the recognition of their incompetence that would ideally motivate them to improve their efforts (Dunning et al., 2004, Kruger and Dunning, 1999).

To make matters worse, the least-skilled individuals tend to disagree with negative feedback (Korsgaard, 1996) and are the most resistant to feedback that would lead to their improvement (Sheldon, Dunning, and Ames, 2014). For example, Sheldon et al. (2014) found that low performers resisted purchasing a book or hiring a coach to assist them with self-improvement, whereas the highest performers voluntarily took these actions, even when they did not substantially benefit by doing so.

The cause of this discrepancy between self-perception and reality is thought to be a lack of monitoring and self-control of one's thinking, a set of processes generally known as *metacognition* (Flavell, 1979, Hacker and Dunlosky, 2003, Sternberg and Kaufman, 1998). Metacognition is often conceptualized as an executive-level control over one's own thinking, including functions such as planning how one will approach a task, monitoring one's success on the task while it is being performed, and evaluating how well one has met the goals of the task after it is completed. Metacognition is sufficiently central to learning that there is an academic journal devoted to it (*Metacognition and Learning*). Effective learners are skilled metacognitive thinkers, and metacognition is fundamental to overall success in life.

The present study examined undergraduate students' responses to feedback on the effort they had invested while solving a series of probability word problems. Mental effort was measured using eye-tracking equipment that provided real-time effort data presented to participants after each problem. Participants were then asked whether the eye-tracking estimate of their workload was too low, too high, or about right. The literature cited above suggests that participants would not likely be in full agreement with the eye-tracking estimates. This study investigates that question, namely, whether undergraduate students solving mathematics problems display flaws in their self-assessment and evaluation of feedback, especially under low-knowledge conditions.

Method

Participants

Fifty-eight undergraduate students (41 women, 17 men) from an eastern U.S. college participated in this experiment. The largest proportion of participants was freshmen (67%), followed by sophomores (21%), juniors (10%), and seniors (2%). Participants reported a mean GPA of 3.14, and most of them (72%) had taken no more than one college mathematics class.

Materials

This experiment was entirely computer-based so that eye-tracking equipment could monitor participants' mental workload as they completed the materials. The computer monitor was a 22in display at a resolution of 1920 × 1080 pixels. The eye-tracking hardware was the 60Hz binocular FOVIO device (2014); the software used to collect and analyze the eye-tracking data was EyeWorks™ version 3.21 (2017). The EyeWorks software provides both raw and analyzed data on pupil size as well as a mental workload measure called the Index of Cognitive Activity (ICA). The ICA algorithm examines sudden changes in pupil size to estimate the amount of mental effort the participant is experiencing during each second of the activity.

The text of the instructional materials described some general probability concepts and then explained total probability, with examples. Two total-probability sample problems were introduced and their solutions were presented and explained. Participants in the diagram condition received instruction on how to use a Venn diagram as an aid when constructing an equation to solve problems involving total probability; participants in the no-diagram condition received instruction on how to generate an equation, without any reference to a diagram.

Following the sample problems were five practice problems; the practice problems were presented in a faded worked-example format (Renkl and Atkinson, 2003), with the first problem fully worked out, the second problem worked out except for the final step, etc., until the fifth problem was left for participants to solve independently with only the steps labeled as prompts. After each of these practice problems, participants were asked how much effort they experienced while completing their solution (see Figure 1). Simultaneously, the EyeWorks software analyzed participants' workload using the ICA algorithm. After participants had submitted their self-ratings, they were presented with the ICA estimate and asked if they agreed with it (see Figure 2).

Figure 1. *Prompt for self-Rating of Effort*

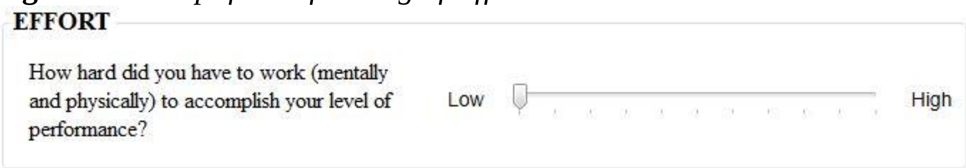


Figure 2. *Prompt for Submitting Agreement with ICA Estimate*

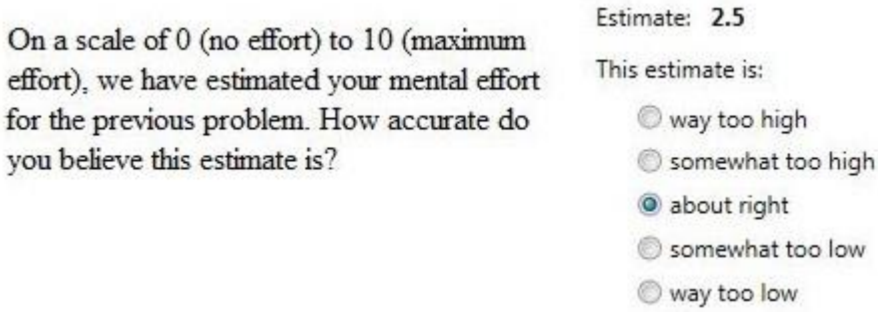
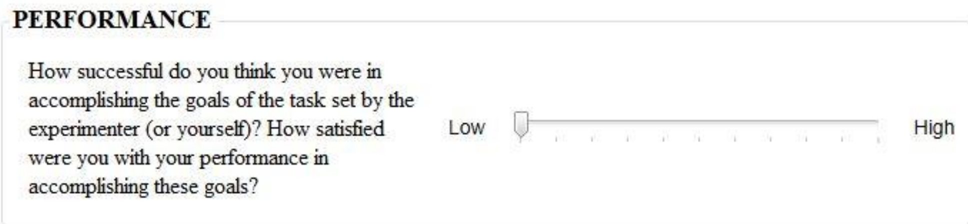


Figure 3. *Prompt for Self-rating of Performance*



The posttest measured participants' ability to solve six total-probability word problems without prompts or other assistance. Participants submitted their mental-demand ratings after each of these problems also, in the same fashion as with the practice problems; ICA data were collected for comparison with the self-ratings; and participants submitted their agreement with the ICA estimates using the same process as with the practice problems. At the conclusion of the posttest, participants were asked a series of questions, including how successful they believed they had been in solving the six posttest problems (see Figure 3).

Design

This experiment used a two-group (diagram vs. no-diagram), between-subjects design. However, for the present paper, these two groups are collapsed for analysis because this manipulation does not pertain to the investigation of metacognitive awareness.

Procedure

Participants completed the experiment individually in a laboratory setting under the direction of an experimenter. All tasks were delivered on a computer. After a brief calibration with the eye-tracking equipment, a demographic survey was administered; then the experiment began with a brief pretest (used to measure participants' prior knowledge of probability concepts, but not relevant for this paper), followed by the instructional materials that included two sample problems and five practice problems. The final portion was the posttest. As previously described, at several points within the experiment, participants were asked to rate how much effort the preceding task had demanded, and how much they agreed with the ICA estimates of their mental workload.

Results

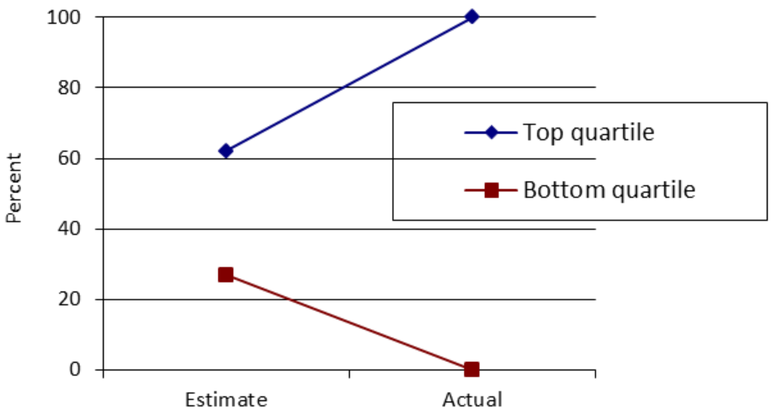
The data in Table 1 show that just over half of the participants judged the ICA estimate to be "about right." The remainder were split on whether the ICA estimate was too high or too low. More participants thought ICA estimate was too low than too high; in other words, many participants believed they had invested more effort in solving these problems than the ICA estimate indicated.

Table 1. *Percentage Agreement with ICA on Posttest Problems*

Problem	Very Low	Low	About Right	High	Very High
1	5%	22%	62%	9%	2%
2	3%	10%	60%	24%	2%
3	9%	10%	62%	14%	5%
4	5%	21%	59%	12%	3%
5	5%	17%	55%	17%	5%
6	2%	16%	60%	14%	9%

The Dunning-Kruger Effect (Dunning et al., 2004, Dunning et al., 2003, Kruger and Dunning, 1999) is documented in Figure 4. As seen on the right-hand portion of the figure, the top quartile of participants achieved a perfect score on the six-item posttest, and the bottom quartile scored zero points. When they were asked to estimate their posttest performance (using the prompt shown in Figure 3), these two subgroups evaluated their success very differently. The top quartile judged themselves to be less successful than they actually were, and the bottom quartile evaluated their performance quite a lot higher than it actually was. This Dunning-Kruger Effect has been documented in the empirical literature in a variety of academic and non-academic domains (Dunning et al., 2003, Ehrlinger et al., 2008, Kruger and Dunning, 1999).

Figure 4. *Estimated and Actual Performance for the Top and Bottom Quartiles on the Posttest*



Discussion

Consistent with expectations indicated by the literature, participants did not uniformly agree with the objective feedback on their effort provided by the eye-tracking measures. Only about half of the participants thought the ICA estimate were about right; many others believed they had worked harder than the estimates indicated. This raises interesting questions for instructors when struggling students approach them claiming they have worked exceptionally hard on a task. These efforts might also be overestimated – yet it would be a mistake for an instructor to make such an assumption even though empirical trends indicate it is likely a valid conclusion.

Also consistent with the literature, the participants with lowest performance in this study had an overly optimistic view of their success (the Dunning-Kruger Effect). This is a potentially paralyzing companion to the foregoing result in which effort is overestimated. Students who believe they have (a) worked harder than they actually have and (b) achieved more than they actually have are not likely to be receptive to corrective feedback that indicates they need to improve their performance.

Conversely, the most successful individuals in this study underestimated their performance, which is also consistent with the empirical literature. Of course, this is not nearly as damaging for forward progress, yet it is important for instructors of such students to ensure they receive positive feedback regarding their progress.

In summary, these data, in concert with the extant literature, document that instructors face a significant challenge when working with lower-performing students. The resistance of these learners to grapple with their true level of performance presents a situation in which they are not likely to succeed – not because they lack the capacity, but because they lack the metacognitive awareness that would contribute to their success. They are saddled with a double curse: they are unable to recognize their weaknesses, and as a result they are ignorant of their need to improve when they are so metacognitively unaware.

References

Beaudoin, M., and Desrichard, O. 2011. Are memory self-efficacy and memory performance related? A meta-analysis. *Psychological Bulletin, 137*, 211-241.

Chemers, M. M., Hu, L.-t., and Garcia, B. F. 2001. Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology, 93*, 55-64.

Dufner, M., Gebauer, J. E., Sedikides, C., and Denissen, J. J. A. 2018. Self-enhancement and psychological adjustment: A meta-analytic review. *Personality and Social Psychology Review*, 1-25.

Dunning, D., Heath, C., and Sula, J. M. 2004. Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*, 69-106.

Dunning, D., Johnson, K., Ehrlinger, J., and Kruger, J. 2003. Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*, 83-87.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., and Kruger, J. 2008. Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*, 98-121.

EyeWorks. Version 3.21 [Computer software]. San Diego, CA: EyeTracking, Inc.

Ferraro, P. J. 2010. Know thyself: Competence and self-awareness. *Atlantic Economic Journal, 38*, 183-196.

Flavell, J. H. 1979. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906-911.

FOVIO. 2014. Canberra, Australia: Seeing Machines.
Freund, P. A., and Kasten, N. 2012. How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin, 138*, 296-321.
Hacker, D. J., and Dunlosky, J. 2003. Not all metacognition is created equal. *New Directions for Teaching and Learning, 95*, 73-79.
Hepper, E. G., Gramzow, R. H., and Sedikides, C. 2010. Individual differences in self-enhancement and self-protection strategies: An integrative analysis. *Journal of Personality, 78*, 781-814.
Kim, Y.-H., Chiu, C.-y., and Zou, Z. 2010. Know thyself: Misperceptions of actual performance undermine achievement motivation, future performance, and subjective well-being. *Journal of Personality and Social Psychology, 99*, 395-409.
Korsgaard, M. A. 1996. The impact of self-appraisals on reactions to feedback from others: The role of self-enhancement and self-consistency concerns. *Journal of Organizational Behavior, 17*, 301-311.
Kruger, J., and Dunning, D. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121-1134.
Lew, M. D. N., Alwis, W. A. M., and Schmidt, H. G. 2010. Accuracy of students' self-assessment and their beliefs about its utility. *Assessment & Evaluation in Higher Education, 35*, 135-156.
Pennycook, G., Ross, R. M., Koehler, D. J., and Fugelsang, J. A. 2017. Dunning-Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin & Review, 24*, 1774-1784.
Renkl, A., and Atkinson, R. K. 2003. Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*, 15-22.
Sheldon, O. J., Dunning, D., and Ames, D. R. 2014. Emotionally unskilled, unaware, and uninterested in learning more: Reactions to feedback about deficits in emotional intelligence. *Journal of Applied Psychology, 99*, 125-137.
Sternberg, R. J., and Kaufman, J. C. 1998. Human abilities. *Annual Review of Psychology, 49*, 479-502.
Vazire, S., and Mehl, M. R. 2008. Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology, 95*, 1202-1216.
Zell, E., and Krizan, Z. 2014. Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science, 9*, 111-125.