

October 2018

How Metacognitive Awareness Relates to Overconfidence in Interval Judgments

Darci Klein

Oklahoma State University, darcik@ostatemail.okstate.edu

Arielle Cunningham

Missouri State University, Arielle924@live.missouristate.edu

Erin M. Buchanan

Missouri State University, erinbuchanan@missouristate.edu

Follow this and additional works at: <https://commons.pacificu.edu/ijurca>

Recommended Citation

Klein, Darci; Cunningham, Arielle; and Buchanan, Erin M. (2018) "How Metacognitive Awareness Relates to Overconfidence in Interval Judgments," *International Journal of Undergraduate Research and Creative Activities*: Vol. 10, Article 6.
DOI: <https://doi.org/10.7710/2168-0620.1109>

How Metacognitive Awareness Relates to Overconfidence in Interval Judgments

Peer Review

This work has undergone a double-blind review by a minimum of two faculty members from institutions of higher learning from around the world. The faculty reviewers have expertise in disciplines closely related to those represented by this work. If possible, the work was also reviewed by undergraduates in collaboration with the faculty reviewers.

Abstract

Making judgments is an important part of everyday life, and overconfidence in these judgments can lead to serious consequences. Two potential factors influencing overconfidence are metacognitive awareness, or the awareness of one's own learning, and the hard-easy effect, which states that overconfidence is more prevalent in difficult tasks while underconfidence is more prevalent in easy tasks. Overall, we hypothesized that participants' metacognitive awareness would significantly relate to their overconfidence levels. Specific hypotheses were that those participants who display higher levels of metacognitive awareness will have lower levels of overconfidence, that harder questions will elicit higher levels of overconfidence and easy questions will elicit underconfidence (congruent with the hard-easy effect), and that the lower range and upper range will on average be equal, with the exact estimate as the midpoint. Participants ($N = 49$) completed a questionnaire containing a set of hard and easy general knowledge questions followed by the Metacognitive Awareness Inventory. The correlation between metacognitive awareness and confidence was negative for hard questions and positive for easy questions. Furthermore, the ranges for easy questions were smaller, resulting in more overconfidence, and the ranges for the hard questions were larger, resulting in underconfidence, thus, showing the opposite of our expected hypotheses.

Keywords

metacognition, psychology, cognition, judgments

INTRODUCTION

Individuals are often overconfident in evaluating the correctness of their knowledge. Over the years, overconfidence has been defined in many ways. Lichtenstein, Fischhoff, and Phillips (1977) describe overconfidence as “the major systematic deviation from perfect calibration... an unwarranted belief in the correctness of one’s answers” (p. 108). Two decades later, Juslin, Winman, and Olsson (2000) term overconfidence as “...the mean subjective probability assigned to the correctness of answers to general knowledge items...tends to exceed the proportion of correct answers” (p. 384). Our daily lives are full of judgment decisions, such as which path to take to quickly get to work, how much time it will take to complete a task, or how much money to save for an upcoming event, but not all estimates can be completely accurate. Soll and Klayman (2004) found that judges who were 80% confident in their decisions were only correct 48% of the time. In their study, participants were asked to answer a numeric general knowledge question, such as “how tall is the Empire State Building?” Given this answer, participants then provided a range around their answer such that they were at least 80% confident that their range included the real answer. Their findings show that estimation often leaves room for a significant amount of error. Consequences can be severe in certain situations involving overconfidence. For example, overestimating the distance between automobiles has the potential to cause an accident.

It has been suggested that overconfidence can result from insufficient cognitive processing (Sniezek, Paese, & Switzer, 1990), information processing biases (Koriat, Lichtenstein, & Fischhoff, 1980), or estimations based solely on

personal experiences (Winman, Hansson, & Juslin, 2004). Sniezek et al. (1990) portrayed that overconfidence was higher when choices were less thought out, and participants should be cued to alternatives for more accurate choices. Koriat et al. (1980) demonstrated that listing contradicting reasons for judgment choice helped decrease overconfidence, and that listing confirmatory reasons (i.e. justifications) for their choice increased overconfidence. Selective retrieval and pulling solely from one’s own experience have also been hypothesized as contributors to bias, which would in turn lead to overconfidence (Winman et al., 2004). In this study, we investigated another possible factor that may increase or decrease overconfidence: metacognitive awareness. Metacognition is “the ability to reflect upon, understand, and control one’s learning” (Schraw & Dennison, 1994, p. 460).

Metacognition has been the topic of a large body of research, usually focused on the poor calibration of subjects’ judgments of their own learning (Koriat & Bjork, 2005; 2006), which can specifically affect studying for student learners (Metcalf & Finn, 2008). This idea of metacognitive awareness involved in overconfidence levels explored in the current study is complimented by a concept called grain size, coined by Yaniv and Foster (1995). Grain size implies that the amount of expertise an individual holds about the topic in question will be represented in their confidence and the way in which their interval answers are represented. For example, “I think the answer is in the hundreds” is a much broader response than “I think the answer is around 300”. According to Yaniv and Foster (1995), the latter answer would imply a greater understanding of the topic compared to the first answer.

Multiple factors have been found to increase overconfidence in judgments, such as the availability of information (Oskamp, 1965) and the level of difficulty (Juslin et al., 2000). The hard-easy effect illustrates these factors and alludes to an effect between over/underconfidence and task difficulty. Overconfidence is more prevalent in difficult items, whereas underconfidence appears more in easy items. Therefore, the more difficult a question, the more likely the responder will limit their answer range, implying a higher level of confidence. The present study considered the hard-easy effect when answering general knowledge questions and analyzing data by varying the difficulty of items to expand upon the hard-easy effect with confidence limits. Not only do overconfidence levels fluctuate based on the difficulty of the question, but information processing is also affected by how the question is presented.

Half-range, full-range, and interval production formats have all been popular designs for general knowledge questions (Winman et al., 2004). Half-range refers to a question in which the participant is presented with a yes or no question and then asked to rate their confidence level (0-100%). A full-range question makes a statement and then participants are asked to rate the probability of that statement being correct (0-100%). Lastly, the interval production format includes a question in which a number is the answer (i.e., What is the population of Japan?), and the participant must give a lower limit and upper limit (1 million to 10 million people). The participant must also assign a level of confidence to their answer (0-100%). Of the three, interval production format creates the highest levels of overconfidence followed by the full-range and half-range formats (Winman et al., 2004).

For the present study, we focused on interval style questions to create the largest opportunity for overconfidence. Overconfidence was assessed using intervals: by subtracting an exact estimate from a lower and upper estimate for each participant. This procedure gave us estimate ranges for each question. Overconfidence resulted when the ranges were small, and underconfidence resulted when ranges were high. For example, if a person estimated that the cost of their car repairs would be \$500 with a range of \$450 – \$550, their confidence range would be 50 points for the upper and lower limit of the car repairs. This range would be overconfident in comparison to a range of \$300 – \$700, which has a confidence range of \$200 (underconfident).

Soll and Klayman (2004) explored multiple interval formats where confidence levels were evaluated, including one-point and two-point formats. One-point format requires the participant to form one statement in which they choose a lower estimate and upper estimate, creating a range where the possible answer lies (“I am 90% sure that this happened between ___ and ___”). Two-point format requires the participant to create two separate statements (I am 80% sure that this happened before _____. I am 80% sure that this happened after ____). In one-point format, participants demonstrated 41% average overconfidence. In two-point format, participants demonstrated lower overconfidence at 23%. Therefore, the two-point format lowered overconfidence levels by creating a range of values that participants were at least 80% or 90% confidence included the exact answer for the general knowledge question. The current study emphasized Soll and Klayman’s two-point format with an added variable. We included an “exact guess” in which the participant made an estimate as to what the true answer was; which in turn created a lower limit range (lower answer to

the exact answer) and an upper limit range (exact answer to the upper limit). For this study, we implemented Soll and Klayman's two-point format because it was comprised of two separate statements, and this encouraged participants to sample their knowledge twice (once for lower and once for upper limits). This format fed into the metacognitive awareness we wished to explore. Presenting multiple opportunities for the participant to examine their experiences and knowledge to answer the questions allowed more chances to exercise their metacognition.

In the present study, we explored the relationship between over/underconfidence levels and the participants' level of specified metacognitive awareness. We used a general knowledge questionnaire, created by the investigators of this study, requiring answers in the form of intervals, as described above. Participants also completed the Metacognitive Awareness Inventory (Schraw & Dennison, 1994). In addition to assessing overconfidence and metacognition, we will be addressing the relationship between overconfidence and question difficulty. Using the interval format allowed us to create a lower range and an upper range (lower estimate to exact estimate and exact estimate to upper estimate), which enabled us to detect any trends in the range ratios.

Overall, we hypothesized that participants' metacognitive awareness would significantly relate to their overconfidence levels. For example, an individual with higher metacognitive awareness would have lower levels of overconfidence because they may not restrict their estimate range, while an individual with lower metacognitive awareness would have higher levels of overconfidence because they may restrict their estimates to smaller ranges. We assumed that those who are more

metacognitively aware would have a better ability to estimate answers with understanding and consideration of the accuracy and restrictions of their knowledge. Three specific hypotheses were constructed for this experiment. First, those participants who display higher levels of metacognitive awareness will have lower levels of overconfidence (assessed by correlation analyses). Second, we expect, congruent with the hard-easy effect, that harder questions will elicit higher levels of overconfidence and easy questions will elicit underconfidence. Lastly, the lower range and upper range will on average be equal, with the exact estimate as the midpoint, which will be assessed with ANOVA on the estimates and confidence ranges.

METHODS

Participants

Forty-nine undergraduate students (nine males and forty females) at Missouri State University participated in this experiment to satisfy a course requirement (Introduction to Psychology, PSY 121). Participants ranged from 18 to 34 years old. The majority of participants were 18 or 19 years old (85.7%), and 35 freshman (71.4%), 11 sophomores (22.4%), and three juniors (6.1%) participated in the experiment.

Measures

Participants completed two measures, a general knowledge questionnaire and the Metacognitive Awareness Inventory (Schraw & Dennison, 1994). The researchers constructed the questionnaire on Qualtrics.com, a survey management site, for the participants to access. It consisted of nine easy general knowledge questions and nine hard general knowledge questions, which are included in Appendix A. This questionnaire was developed by searching for and creating

trivia style questions that an average person might be able to answer. They were sorted into easy and hard questions based on feedback from pilot testing. Twenty-five undergraduate students at Missouri State University pilot tested the 30 general knowledge questions, prior to the experiment, to confirm difficulty level. The percent correct for each question was calculated and the nine hardest and nine easiest questions were included in the experiment questionnaire. Each question was designed to solicit a number as the answer, with the ability to create a range of estimates around that number. Researchers considered the answers correct if they were within three numbers above or below the correct answer. The final survey questions can be found in Appendix A. The entire survey can be found at <https://osf.io/ept3c/>, along with the IRB, data, and analysis files.

Participants provided a range of answers to each general knowledge question, a lower limit, exact estimate, and upper limit. For example, the participant was asked, "How many feet apart are major league baseball bases?" The answer is 90 feet. Therefore, the student stated the lowest possible answer they thought it could be, possibly 50 feet (lower limit), an estimate of the true answer, possibly 90 feet (exact estimate), and the highest possible answer they thought it could be, possibly 130 feet (upper limit). These questions were ordered randomly for each participant to account for ordering effects.

After completing the general knowledge questionnaire, the participants completed the Metacognitive Awareness Inventory (MAI). The MAI is a 52-question self-report survey that assesses the ways one strategizes, thinks, and understands their learning. These questions were presented as statements such as, "I draw pictures or diagrams to help me understand while learning." The MAI presents each self-report

item as a True/False statement. Participants completed the 52 statements in the order intended by the original authors. The 52 statements comprised two subcategories: knowledge of cognition and regulation of cognition. Knowledge of cognition refers to participants' knowledge about themselves, strategies, and situations when said strategies are useful. Examples of knowledge of cognition statements in the MAI are "I understand my intellectual strengths and weaknesses" and "I try to use strategies that have worked in the past" (Schraw & Dennison 1994). Regulation of cognition relates to students' understanding of the way they plan, monitor, and evaluate their learning, and the way they apply strategies. Examples of regulation of cognition statements in the MAI are: "I ask myself questions about the material before I begin" and "I change strategies when I fail to understand".

Schraw and Dennison (1994) found this two-component model to be valid. The internal consistency of these scales ranged from .88 to .93. Schraw and Dennison also found that while knowledge and regulation of cognition were represented in the MAI, both of these components function independently of one another, each making a unique contribution to cognition.

Procedure

Participants registered for this experiment through Missouri State University's SONA-system. The experiment was held in a computer lab, implemented on a standard PC, and it lasted between 10 and 25 minutes (depending on participants' speed of completion). Once participants registered for a timeslot, they arrived at the computer lab and the primary investigator prompted the students. The investigator requested the participants refrain from using electronic devices, such as their cellular phones, during the experiment, to prevent

the students from looking up the answers to the general knowledge questions. The investigator wrote the URL to the Qualtrics survey on the board at the front of the room, and the participants entered the web address in an internet browser.

Upon opening the survey, a consent form was presented. If participants selected “Yes”, the survey proceeded to the 18 general knowledge questions and the Metacognitive Awareness Inventory (MAI), but if the student chose “No,” they were directed to an end-of-survey prompt and thanked for their time. For the students who continued with the survey after the consent form, the 18 general knowledge questions were randomly presented one at a time. Students were asked to enter a lower-limit, exact answer, and upper-limit as described in Materials. After completion of the general knowledge questionnaire, the students were asked to complete the 52-item self-report MAI.

Students who completed the experiment were asked to follow a link to an independent survey on Qualtrics and enter their first and last names. The names on the independent survey allowed the investigators to give credit to the students for their Introduction to Psychology (PSY 121) course without compromising their anonymity. Students were emailed debriefing information following the experiment.

RESULTS

Data Processing

All estimates were recoded to standardized estimates across questions by dividing each estimate (lower, exact, upper) by the correct answer. Therefore, if a participant’s estimate was equivalent to the correct answer, their standardized estimate would equal 1. Estimates below the correct answer were less than 1, and estimates

higher than the correct answer were greater than 1. Next, the investigators created confidence levels for the lower estimates and upper estimates by subtracting the exact estimate from each one. This procedure created estimate ranges where low scores indicated overconfidence and high scores indicated underconfidence. Overconfidence was found in narrow ranges, implying the participant limited their answers. Conversely, underconfidence was found in broad ranges, implying the participants created ranges excessively beyond the scope of possible correct answers.

Extreme estimates were eliminated, as identified by examining a histogram of the standardized estimate scores. For instance, 6500 as a standardized estimate was not used in analysis. Average scores for each participant for estimates, (lower, exact, upper) and confidence levels (lower, upper ranges) were created by averaging across easy and hard questions separately.

The MAI was scored by totaling the questions for knowledge and regulation of cognition separately according to Schraw and Dennison’s scoring guidelines. See Table 1 for means, standard deviations, minimums, and maximums.

EASY	Min	Max	Mean	SD
Lower Estimate	0.01	0.61	0.20	0.13
Exact Estimate	0.52	1.17	0.92	0.15
Upper Estimate	0.01	0.53	0.16	0.11
Lower Range	0.40	1.08	0.75	0.15
Upper Range	0.55	1.37	1.03	0.17
HARD	Min	Max	Mean	SD
Lower Estimate				
Exact Estimate	0.37	1.30	0.77	0.20
Upper Estimate	0.03	0.69	0.25	0.15
Lower Range	0.28	0.99	0.59	0.19
Upper Range	0.41	1.34	0.86	0.26
Knowledge about Cognition	9	17	13.94	2.23
Regulation of Cognition	13	34	26.84	4.87

Table 1. Minimums, maximums, means, and standard deviations of easy and hard (lower, exact, upper estimates and lower/upper ranges) and knowledge about cognition and regulation of cognition.

Correlation of Confidence with Metacognition

When analyzing metacognitive awareness and its predictive ability to estimate overconfidence levels, no significant correlation was found between MAI subscales and confidence levels (see Table 2 for r and p -values). Interestingly, metacognitive awareness for hard questions resulted in negative correlations while easy questions resulted in positive correlations. This finding shows that, for easy questions, participants whose metacognitive awareness was higher also had larger ranges, implying underconfidence. On hard questions, those who had higher levels of metacognitive awareness had smaller ranges, implying overconfidence. This result was the opposite of what we had originally hypothesized.

EASY	Knowledge about Cognition	Regulation of Cognition
Lower Limit	.041, $p = .780$.176, $p = .227$
Upper Limit	.111, $p = .448$.180, $p = .215$
HARD		
Lower Limit	-.150, $p = .305$	-.221, $p = .127$
Upper Limit	-.101, $p = .491$	-.146, $p = .317$

Table 2. Correlation of question difficulty, upper and lower limits, and subscales of MAI. $df = 47$.

Average Correct Scores

A 2 (easy, hard questions) \times 3 (lower, exact, upper estimate) Repeated Measures ANOVA addressed the correctness of the participants' answers. We found difficulty of the question was paralleled by the number of correct answers, $F(1, 48) = 29.09$, $p < .001$, $n_p^2 = .38$. Therefore, participants answered easy questions ($M = 0.90$, $SE = 0.02$) correctly more often than hard questions ($M = 0.74$, $SE = 0.03$), as defined by being closer to 1 on their standardized estimates. A significant effect of correctness for the lower, exact, and upper estimates was found, $F(1, 48) = 116.09$, $p < .001$, $n_p^2 = .71$. The investigators analyzed this effect by using dependent t -tests to examine the differences between lower-exact and exact-upper ranges to determine the change in estimate across questions. The range from lower ($M = 0.67$, $SE = 0.13$) to exact ($M = 0.84$, $SE = 0.13$) differed significantly, $t(48) = -10.48$, $p < .001$, $d = -1.50$, as did the range from exact to upper ($M = 0.95$, $SE = .18$), $t(48) = -8.15$, $p < .001$, $d = -1.16$.

This finding implies that the confidence estimates are different across questions and that there was a real and varied range around the exact estimate. Therefore, the lower, exact, and upper estimates were not repeatedly reported as the same number or in set patterns (i.e. participants did not enter 3, 3, 3, or 10, 11,

12, as lower, exact, and upper estimates consistently across questions). This result shows that our manipulation was successful, and we were able to analyze over- or underconfidence because participants' answers varied from the exact estimate. See Figure 1 for the relationship between question difficulty and average standardized estimate.

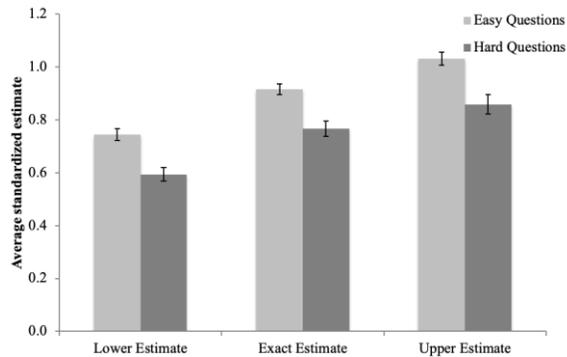


Figure 1. Relationship between average standardized lower, exact, and upper estimates and question difficulty.

Confidence Ranges

To analyze the size of lower ranges and upper ranges with easy and hard questions, we ran a 2 (easy, hard questions) x 2 (lower, upper range) Repeated Measures ANOVA. The main effect of easy versus hard questions was significant, $F(1, 48) = 7.86, p = .007, \eta_p^2 = .14$. According to this analysis, the ranges for easy questions ($M = 0.18, SE = 0.02$) were smaller, resulting in more overconfidence, and the ranges for the hard questions ($M = 0.24, SE = 0.02$) were larger, resulting in underconfidence. This result was contrary to our hypothesis. We believed that the participants would demonstrate consistency with the hard-easy effect in that hard questions would elicit overconfidence and easy questions would elicit underconfidence, but this hypothesis was not supported. The main effect of confidence range was also significant, $F(1, 48) = 6.05, p = .018, \eta_p^2 = .11$. Lower confidence ranges overall were larger ($M =$

0.23, $SE = 0.03$) than upper confidence ranges ($M = 0.19, SE = 0.02$). Again, the interaction between range confidence size and question difficulty was not significant, $F(1, 48) = .31, p = .580, \eta_p^2 = .01$. Figure 2 shows the means for each.

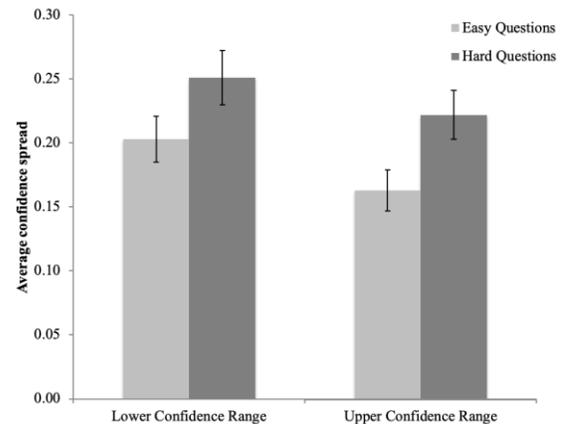


Figure 2. Relationship between the average spread of lower and upper confidence ranges and question difficulty

DISCUSSION

Contrary to our hypothesis, we found that the Metacognitive Awareness Inventory subscales were not related to participants' level of overconfidence on interval-style general knowledge questions. However, metacognitive awareness for hard questions resulted in a negative correlation, which shows that individuals with higher scores on the MAI also had smaller ranges, implying overconfidence. Furthermore, individuals who scored high on the MAI had larger ranges for easy questions, resulting in underconfidence. This result is the opposite of what we hypothesized. We assumed that when individuals were more aware of their thinking processes, they would create larger ranges to encompass all possible answers, but in reality the participants with higher levels of reported metacognition created smaller ranges to questions. Therefore, general overconfidence shown by

individuals may be more enhanced when an individual believes they have more metacognitive skills than others. Potentially, this result is related to the Dunning-Kruger effect in that people are often overconfident in their abilities, even when performance is low (Kruger & Dunning, 1999).

Koriat and Bjork (2005, 2006) showed similar results when examining judgments of learning, such that people are particularly poor at estimating their skills on a future test without specific study-test practice. Metcalfe and Finn (2008) have shown the importance of this research to student learners, as the overconfidence in judgments influenced their choices to continue to study. Therefore, as a person learns information, they also judge the strength of that learned information. If these judgments are overconfident, the person may discontinue studying, which could have dire consequences for course completion in primary and secondary education. Age may additionally play a role in these results, as our study contained predominately college freshman, who may not have had the practice and feedback at these types of estimations that can help tune them more accurately (England & Serra, 2012).

Question difficulty was paralleled by the number of correct answers. Therefore, easy questions were answered correctly more often than hard questions. This finding was expected based on pilot testing. The main effect of easy versus hard questions was significant. Easy questions had smaller ranges, resulting in overconfidence, and the ranges for the hard questions were larger, resulting in underconfidence. This result was contrary to the hard-easy effect, which we thought would apply in this study; however, our results can likely be tied to the “better-than-average” effect in which the hard-easy effect can reverse with task difficulty and overconfidence (Larrick, Kurson, & Soll, 2007). The better-than-average effect is

often found when participants overestimate their abilities to perform specific tasks, and motivated reasoning (i.e., wanting to view oneself in a positive light) has been proposed as a likely reason for this effect (Taylor & Brown, 1998). These effects are important to understand because of their relationship to complex decision making, such as business takeovers (Camerer & Lovallo, 1999) and employment strikes (Babcock & Olson, 1992).

In future research, we recommend acquiring a larger sample size. This study had a limited participant pool and a short time window to collect data. Participants may have been potentially uninterested in the study because it was required for course credit, resulting in less accurate self-report results. Our student population is largely female, thus driving the larger number of female participants. Our results appear to represent female students, and further studies may wish to investigate if male students show a different pattern of results with a larger subsample size. This study and others in the future can enhance our understanding of decision making and what factors affect it. The idea that high metacognitive awareness would decrease the amount of overconfidence in interval style questions remains a strong argument in our minds, especially as it relates to testing environments that students might encounter. Increasing people’s awareness of the way they analyze questions could allow them to recognize that the range may be broader to include possible answers in harder questions. Creating a novel metacognitive awareness assessment tool could be an option for finding more reliable results.

REFERENCES

- Babcock, L., & Olson, C. (1992). The causes of impasses in labor disputes. *Industrial Relations*, 31, 348-360.

- doi: 10.1111/j.1468-232x.1992.tb00313.x
- Camerer, C. F., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89, 306-318. doi: 10.1257/aer.89.1.306
- England, B. D. & Serra, M. J. (2012). The contributions of anchoring and past-test performance to the underconfidence-with-practice effect. *Psychonomic Bulletin & Review*, 19(4), 715-722. doi: 10.3758/s13423-012-0237-7
- Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard–easy effect. *Psychological Review*, 107(2), 384-396. doi: 10.1037/0033-295x.107.2.384
- Koriat, A. & Bjork, R.A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2), 187-194. doi: 10.1037/0278-7393.31.2.187
- Koriat, A. & Bjork, R.A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition*, 34(5), 959-972. doi: 10.3758/bf03193244
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal Of Experimental Psychology: Human Learning And Memory*, 6(2), 107-118. doi: 10.1037/0278-7393.6.2.107
- Kruger, J. & 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(6), 1121-1134. doi: 10.1037/0022-3514.77.6.1121
- Larrick, R. P., Burson, K. A., & Soll, J. B. (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes*, 102(1), 76–94. doi: 10.1016/j.obhdp.2006.10.002
- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1977). Calibration of probabilities: The state of the art. In H. Jungermann & G. de Zeeuw (Eds.), *Decision making and change in human affairs*. Amsterdam. doi: 10.1007/978-94-010-1276-8_19
- Metcalf, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15(1), 174-179. doi: 10.3758/PBR.15.1.174
- Oskamp S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, 29, 261–265. doi: 10.1017/cbo9780511809477.021
- Schraw, G., & Dennison, R. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460-475. doi: 10.1006/ceps.1994.1033
- Snizek, J. A., Paese, P. W., & Switzer, F. S. (1990). The effect of choosing on confidence in choice. *Organizational Behavior And Human Decision Processes*, 46(2), 264-282. doi: 10.1016/0749-5978(90)90032-5
- Soll, J. B., & Klayman, J. (2004). Overconfidence in interval estimates. *Journal Of Experimental*

- Psychology: Learning, Memory, And Cognition*, 30(2), 299-314. doi: 10.1037/0278-7393.30.2.29
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193-210. doi: 10.1037/0033-2909.103.2.193
- Winman, A., Hansson, P., & Juslin, P. (2004). Subjective probability intervals: How to reduce overconfidence by interval evaluation. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*, 30(6), 1167-1175. doi: 10.1037/0278-7393.30.6.1167
- Yaniv, I., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy informativeness trade-off. *Journal of Experimental Psychology: General*, 124, 424-432. doi: 10.1037/0096-3445.124.4.424