Proposed Journal section: Cognitive/Neuroscience

Neurophysiological Correlates of the Dunning-Kruger Effect Reveal Contributions of Episodic Memory to Metacognitive Judgments of Illusory Superiority

Alana Muller[1,2], Lindsey Sirianni[1], & Richard J. Addante[1]

[1]California State University – San Bernardino, 5500 University Parkway, San Bernardino, CA 92407

[2]University of Arizona, 1423 E University Blvd, Bldg. 67, Rm 223, Tucson, AZ 85719

Corresponding Author:
Richard J. Addante, PhD
California State University – San Bernardino
Department of Psychology
5500 University Parkway
San Bernardino, CA 92407
Richard.Addante@csusb.edu

Running title: Physiological Correlates of Dunning-Kruger Effect

Total number of words: manuscript (11, 804 words); the Abstract (250 words)

Keywords: Metacognition, EEG, ERP, recollection, familiarity

Abstract.

The Dunning-Kruger Effect is a metacognitive phenomenon of illusory superiority in which individuals who perform poorly on a task believe they performed well, yet individuals who performed very well believe they under-performed. This phenomenon has yet to be directly explored in episodic memory, nor explored for reaction times and physiological correlates. We designed a novel method to elicit the Dunning-Kruger Effect (DKE) via a test of item and source recognition while electroencephalography (EEG) was recorded; throughout the task, participants were asked to estimate the percentile in which they performed compared to others. Results revealed participants in the bottom 25th percentile overestimated their percentile the most, while participants in the top 75th percentile underestimated their percentile, exhibiting the classic DKE. Reaction time measures revealed a condition x group interaction whereby over-estimators responded faster than under-estimators when estimating being in the top percentile and responded slower when estimating being in the bottom percentile. Between-group EEG differences were evident between over-estimators and under-estimators when collapsing across all Dunning-Kruger responses, which revealed FN400-like effects of familiarity supporting differences for over-estimators from 400-600ms, whereas 'old-new' memory ERP effects revealed a late parietal component (LPC) associated with recollection-based processing for under-estimators that was not evident for over-estimators. Findings suggest over- and under-estimators use differing cognitive processes when assessing their performance, such that under-estimators rely on recollection during memory and over-estimators draw upon excess familiarity when over-estimating their performance. Episodic memory thus appears to play a contributory role in metacognitive judgments of illusory superiority in the DKE.

Introduction

"*The fool doth think he is wise, but the wise man knows himself to be a fool*"

(Shakespeare, 1601)

The Dunning-Kruger Effect describes the phenomenon in which poor performers on a task tend to overestimate their performance while high performers on a task tend to underestimate their performance. Overconfidence has been a topic of interest throughout recorded history, as early as the time of Socrates, who was noted by Plato for identifying *"…it is likely that neither of us knows anything worthwhile, but he thinks he knows something when he does not, whereas when I do not know, neither do I think I know; so I am likely to be wiser than he to this small extent, that I do not think I know what I do not know"* , and later with Charles Darwin noting more simply that "*Ignorance more frequently begets confidence than does knowledge*" (Darwin, 2009/1871). Also implied by these timeless observations is that these metacognitive illusions are bi-directional, such that the more competent individuals who perform highest tend to also under-estimate their abilities (for Review see Zell et al., 2019).

Perceptions of inaccurate confidence in one's abilities is a common phenomenon that can happen to anyone (including the authors and readers) and can lead to serious problems that are often preventable. For example, the overconfidence of a Captain that the Titanic was unsinkable led to loss of over 1500 lives (Bartlett, 2012; Lord, 1986; Lord, 1955). Conversely, when the most competent people hold back contributions from teams or society because they think others are better-suited, the effects can be a substantial loss on society: we could lose the most competent people for leadership

positions, and instead embrace those whose merely think they are the best (incorrectly). Therefore, it is important to understand how & why these inaccurate judgments of one's abilities occur, so that they can be prevented.

In 1999, the study of over- and under-confidence was further characterized by David Dunning and Justin Kruger, who explored combining the two effects under one term. In their landmark study, Dunning & Kruger conducted several studies showing that bottom performers on a logical reasoning task overestimated their task performance scores and that, conversely, top performers underestimated their task performance scores. The name "The Dunning-Kruger Effect" (DKE) became highly popularized throughout mainstream culture and society (Herrera, 2018; Marczyk, 2017).

Generically, the DKE describes a phenomenon in which self-estimates of performance on a task and the percentile ranking among others also participating in the task do not match performance accuracy and actual rank respectively. The direction of this mis-match of self-perception extends in both directions (Sieber, 1979). Empirically, this paradigm has been used successfully in many different tasks to elicit the DKE on such tasks as knowledge of microeconomics material on a midterm and final (Ryvkin, Krajč, & Ortmann, 2012b), knowledge about the University of Chicago (Burson et al., 2006), ability to identify humorous jokes (Kruger & Dunning, 1999), logical reasoning (Schlösser et al., 2013), cognitive reflection (Pennycook et al., 2017), size judgments (Sanchez, 2016), finance (Atir, Rosenzweig, & Dunning, 2015), and computer programming (Critcher & Dunning, 2009). More broadly, the effect has been referred to in popular culture contexts of aviation, (Woodbury, 2018), driving (Svenson, 1981), and professors rating their own teaching skills (Cross, 1977). However, an account of the

cognitive processes leading to these illusory experiences have yet to be fully explored or understood.

The Dunning-Kruger Effect. The DKE is a psychological phenomenon described by a mismatch in one's perceived ability and the reality of one's objective performance on a given task, and this appears to be directionally moderated by the factor of ability. Low performers (individuals who do not earn high scores on a test using an objective scale) tend to overestimate their performance on a task while high performers (individuals who earn high scores measured on an objective scale) tend to underestimate their performance on the same task. This miscalibration is most often measured using two different questions. The first type of question asks participants to estimate their score using an objective scale (objective performance estimate). The second type of question tends to ask participants to estimate the percentile in which they rank in relation to other students/group of individuals participating in the experiment (relative performance estimate). Researchers of the DKE generally find that low performers tend to overestimate their objective performance on a task, which then inflates their subsequent relative performance estimate (Adams & Adams, 1960; Burson, Larrick, & Klayman, 2006; Ehrlinger & Dunning, 2003; Kruger & Dunning, 1999; Oskamp, 1965; Pennycook, Ross, Koehler, & Fugelsang, 2017; Ryvkin, Krajč, & Ortmann, 2012a; C. Sanchez & Dunning, 2017).

However, there are some different findings about high performers. Dunning and Kruger (1999) found that high performers tend to accurately judge their objective score on a task but underestimate their relative performance score. They argue that low performers' and high performers' scores on their objective score should be rather

different. Even though low performers judge their raw score to be higher than it is, their estimates are not as high as the high performers' estimates. Because high performers perform much better on the task than low performers and they tend to estimate their score accurately, their estimates are above even the inflated estimates of the low performers. However, other studies have found that high performers still underestimate their objective score rather than gauging their score accurately (Burson et al., 2006; Pennycook et al., 2017; Schlösser, Dunning, Johnson, & Kruger, 2013) which does not fit Dunning and Kruger's explanation for high performers' metacognitive errors. Currently, this discrepancy in accounts has not been resolved in the literature.

Most of the paradigms used to research the DKE follow a similar format: participants are given a task such as a series of logical reasoning problems or math problems, etc., and after they finish the task in its entirety, they are asked to estimate their overall objective score on the task itself and/or their relative performance estimate. That is, the data point for their metacognitive judgment is a single data point assessed at the conclusion of the study and it represents their aggregated assessment of performance across a great many of trial instances. That approach has not, however, permitted the ability to measure multiple instances of the cognitive phenomenon per person, and precluded being able to collect simple statistical measures of central tendency, such as reaction times. Accordingly, extant approaches have also precluded the collection of neuroscientific measures that rely upon have multiple measures per person (EEG, fMRI, etc.).

Though relatively limited, researchers have investigated illusory over-confidence for decades. One of the earliest studies of overconfidence was conducted by Adams &

Adams (1960) who found that participants' confidence in their ability to recognize correctly spelled words was higher than their accuracy at the task. Five years later, Oskamp (1965) found that when clinical psychologists were asked to make a diagnosis for a case study, their confidence in their decision increased when they were given more information about the case although their accuracy did not increase. These instances showed that confidence and accuracy were not necessarily correlated in both experimental studies and in more practical issues of clinical diagnoses – a finding which has persisted in modern research on memory as well (Hirst et al., 2015; Kvavilashvili, Mirani, Schlagman, Foley, & Kornbrot, 2009), which will be discussed later below.

Memory research intersects with the DKE at the point of confidence in one's memories and the accuracy of those memories. A large collection of research is available that supports the finding that high confidence does not beget high accuracy (Chua, Hannula, & Ranganath, 2012; Roediger & DeSoto, 2014; Koriat et al., 2008; Nelson & Narens, 1990; Wells, et al., 2002); that is, people can often exhibit illusory over-confidence in their memory judgments, and this has had wide-ranging impacts on society. For example, some of the most salient memories, so-called "flashbulb memories", are found to be no more accurate than other memories despite the participants' high confidence in their accuracy (Neisser & Harsch, 1992; Brown & Kulik, 1977). Other experiments studying the 'false fame' phenomenon demonstrate that familiarity with names can lead to falsely recognizing them as famous later (Dywan & Jacoby, 1990; Jacoby, Kelley, Brown, & Jasechko, 1989; Jacoby, Woloshyn, & Kelley, 1989, 2004): familiarity caused participants to falsely believe an ordinary name was famous because they could not recollect the context in which the name was presented.

Perhaps some of the most impactful research on memory failing to correlate with accuracy pertains to the criminal justice system (Heaton-Armstrong, Shepherd, Gudjonsson, & Wolchover, 2006; Loftus, 1975; Loftus & Zanni, 1975; Nadel & Sinnott-Armstrong, 2012; Pena, Klemfuss, Loftus, & Mindthoff, 2017; Schacter & Loftus, 2013). In a study by Pena et al. (2017) researchers asked participants to make judgments about their accuracy on a memory test for a mock crime observed earlier in a study. They found that participants who performed poorly on the memory test for details of a mock crime overestimated their memory accuracy. These results were consistent with the results of poor performers exhibiting the DKE, demonstrating the link that may exist inherently between the two domains of memory and illusory superiority – yet which remains largely unexplored.

Throughout the years, overconfidence continued to be studied in many different contexts such as social situations (Dunning, Griffin, Milojkovic, & Ross, 1990; Vallone, Griffin, Lin, & Ross, 1990), tasks of differing degrees of difficulty (Bradley, 1981; Lichtenstein & Fischhoff, 1977; Sen & Boe, 1991), and ways to reduce overconfidence (Arkes, Christensen, Lai, & Blumer, 1987; Zechmeister, Rusch, & Markell, 1986). In this research, there was a common finding of overconfidence in wrong answers (Fischhoff, Slovic, & Lichtenstein, 1977; Harvey, 1990; Howell, 1971; Lichtenstein & Fischhoff, 1977; May, 1986) and a less common finding of under-confident correct answers for top performers (Sieber, 1979) since the focus of the research at that time was not the high performers. The term “overconfidence effect” developed to describe this pattern of higher self-estimates of confidence than ability. Throughout the extant research, though,

theoretical frameworks to account for these metacognitive illusions remains relatively sparse.

The Dunning-Kruger Effect is similarly related to other social psychological phenomena such as the 'better-than-average-effect' (BTAE) whereby people are found to rate themselves as better than average peers (for Review see Zell et al., 2019; Alicke & Govorun, 2005; Chambers & Windschitl, 2004; Moore & Healy, 2008; Sedikides, Gaertner, & Cai, 2015; Brown, 1986, 2010), reflecting a general self-evaluation bias (Kwan et al., 2004; Moore & Healy, 2008); though as noted by Zell et al. the DKE emphasizes more of the self-perceptions on a given test/task relative to objective scores on that task as opposed to the BTAE's alternate focus on more generic comparisons to the presumed-average population. Several somewhat cognitive accounts that have been put forth, such as egocentrism, positing that people know more about themselves than about others and hence rank that higher in evaluations from the richer depth of knowledge (Moore & Healy, 2008), but as noted by Zell et al., (2019) such accounts also have not been able to fully explain the underlying mechanisms of what is driving these kinds of social psychological effects of over-confidence in one's abilities.

Theoretical Accounts and Models of the Dunning-Kruger Effect.

Dunning and Kruger postulated that the reason for low performer's incorrect estimation for objective performance scores is due to meta-ignorance or two-fold ignorance (Kruger & Dunning, 1999). This means that poor performers are unaware that they are ignorant of the details needed to correctly complete the task and that double ignorance bolsters feelings of false superiority (Schacter, 2012). More simply, poor

performers do not have the knowledge to complete the task correctly and because they do not know their answers are incorrect, and they hence believe they are performing well (i.e. ‘ignorance is bliss’). For example, poor performers on a task of logical reasoning ability did not have the necessary knowledge to answer the questions in the test correctly. They were also unaware that their answers were incorrect providing them with false confidence that they answered correctly (Schlösser et al., 2013). While this is a very useful behavioral description, it does little to advance an understanding of the cognitive processes involved in the pervasive illusion.

Dunning and Kruger also used what they coined “reach-around-knowledge” to explain low performers’ high confidence in their abilities. The term ‘reach-around knowledge’ refers to a person’s unique knowledge gained from previously participating in a similar task to the task presented and generalizing their past experiences to the current experience (Dunning, 2011). Kruger and Dunning postulated that participants use reach-around knowledge to help achieve their estimation, though this doesn’t necessarily require that it leads them to an accurate perception. According to this view, in order to give an overestimation, one must first have knowledge about the same or similar tasks but not have the knowledge about the details of the task to complete it correctly. Having a larger store of reach-around knowledge should therefore increase the overestimation of poor performer’s scores. On the contrary, having a smaller store of reach-around knowledge should decrease the overestimation of one’s abilities resulting in a more accurate performance estimate.

Dunning and Kruger’s “reach-around-knowledge” account, however, is only a theoretical concept that has not yet been operationally defined or objectively measured,

and also lacks a substantive construct grounded in cognitive psychology. Nevertheless, it provides a useful platform from which to expand upon in investigating this phenomenon within a theoretical construct. The reach-around-knowledge account provided by Dunning and Kruger refers to changes in current behavior based upon prior experience, which is fundamentally a defining feature of memory (Rudy, 2013), and as such it recognizes a key role that memory processes may play in contributing to this metacognitive illusion. There happens to be a rich and robust empirical history of memory processes being both theoretically and operationally defined and studied. Two of these cognitive processes of episodic memory that may contribute to the DKE are familiarity and recollection (Yonelinas, 1999; 2002; Yonelinas et al., 2010). These processes align closely with the general concepts that Dunning and Kruger attributed to their reach-around-knowledge account, and importantly, we can draw upon this platform of memory processes in approaching the DKE in a systematic manner, which will be discussed in depth in the sections below.

Familiarity and Recollection. The cognitive processes of recollection and familiarity have featured prominently in theoretical models of episodic memory for several decades (Eichenbaum, Yonelinas, & Ranganath, 2007; Khoe, Kroll, Yonelinas, Dobbins, & Knight, 2000; Rugg & Curran, 2007; Squire, Wixted, & Clark, 2007; Yonelinas, 1999, 2002). Broadly, familiarity refers to having exposure to some material but not being able to recall the context in which it was presented, usually associated with mid-levels of confidence, whereas recollection refers instead to information that one can recall specific contextual details about from prior episodic experiences and is typically associated only with high-confidence memories (Addante, et al., 2012a,

2012b). In explicit memory, theoretical models of recognition are largely governed by the dual processes of familiarity and recollection (Diana, Yonelinas, & Ranganath, 2008; Eichenbaum et al., 2007; Ranganath, 2010; Yonelinas, 2002; Yonelinas, Aly, Wang, & Koen, 2010) (though see Wixted, 2007 and Wixted & Mickes, 2010 for alternative views), and it is possible that understanding of familiarity and recollection processes in memory may help explain a proportion of variance in the DKE.

Recollection is typically operationalized as the declarative retrieval of episodic information of both the item and context bound together into a cohesive retrieval of the episodic event (for review see Diana et al., 2008), and is usually associated with the retrieval of contextual information surrounding the item of the event (Addante et al. 2012a; for reviews see Eichenbaum et al., 2007; Yonelinas et al. 2010; Ranganath 2010). The item in the event, however, may be retrieved without recollection and via reliance upon familiarity, typically conceptualized as retrieval of an item from a prior episode but without the associated contextual information in which it occurred. Familiarity occurs, for instance, when a person can remember that someone seems familiar from the past but cannot retrieve who the person is or from where they know them. Recollection, on the other hand, would be remembering precisely who someone is and how you know them from a prior episode of one's past experience.

Each of these two memory phenomena have been found to be dissociable cognitive processes (Yonelinas, 2002), with dissociable neural substrates in the medial temporal lobes (Ranganath et al., 2004), neuropsychologically dissociable among patient impairments (Addante, Ranganath, Olichney, & Yonelinas, 2012; Düzel et al., 1999; Mecklinger, von Cramon, & Matthes-von Cramon, 1998; Bowles et al., 2008), and

with distinct patterns of electrophysiology at the scalp that is both spatially and temporally dissociable in event-related potentials (ERPs) (Addante et al., 2012; Curran, 2000; Friedman, 2013; Gherman & Philiastides, 2015; Rugg et al., 1998; Rugg & Curran, 2007). Physiologically, familiarity has been associated with ERP differences in old and new memory trials during a negative-going peak at the mid-frontal scalp sites at approximately 400 milliseconds to 600 milliseconds post stimulus, called the mid-frontal old-new effect, or FN400 (for frontal-N400 effect). On the other hand, recollection has been associated with differences between memory conditions occurring at a peak in the ERP at the parietal region of the scalp from approximately 600 milliseconds to 900 milliseconds, or LPC (Addante et al., 2012; Leynes et al., 2005; for reviews see Rugg & Curran, 2007; Friedman, 2013).

A Memory-Based Framework of the Dunning-Kruger Effect.

Many of the accounts of the DKE have focused primarily upon interpretations based upon metacognition and competency (Adams & Adams, 1960; Ehrlinger & Dunning, 2003; Kruger & Dunning, 1999; Oskamp, 1965; Pennycook et al., 2017; Ryvkin et al., 2012a; C. Sanchez & Dunning, 2017). However, it is very likely that DKEs could also be contributed to by memory experiences in one's past influencing the real-time processing of the current information- either via explicit or implicit means. Based upon the converging literatures from memory and metacognition, a viable alternative theory we postulate here to explain the DKE is that the illusory superiority experience may also be driven, at least in part, by increased familiarity from prior experience of one's past with the tested materials (e.g.: Chua et al., 2012), by which people may use

a decision heuristic deriving supposed high performance despite a lack of specific retrieval of the relevant details that would be involved in marking competency with the material. In this view, the experience of one's lacking in distinct recollection but being generally familiar with material will lead people to assume that they are competent and successful with it and would be associated with increased FN400 amplitudes in ERPs for inaccurate over-estimators.

In that case, for example, it would be a relatively 'dangerous' combination to have insufficient recollection but excessive familiarity with a given topic, stimuli, or information (Chua et al., 2012) because it could lead to inaccurate over-estimates of one's abilities & competencies. Accordingly, under-estimators of self-performance may be marked by having had sufficient recollection of the study material (e.g. competency) such that these instances are associated with an LPC, while also leading people to perhaps recollect the extent of non-criterial information that their cognizance acknowledges could still be relatively wrong, hence lowering their relative scores to other supposed people. In this case, the excess of recollection signal would outweigh the relative noise of unsure familiarity.

Current Study. The current paradigm has been designed to study the metacognitive decision-making process as it occurs in real-time during DKE relative performance estimates that are provided by participants throughout an item recognition memory test. Based upon the memory framework for making metacognitive decisions noted above, broadly, we hypothesized that low performers who tend to overestimate their ability/scores may do so because of familiarity for previous experiences in similar situations; and that high performers who tend to underestimate their ability/score may

use more recollection to accurately outperform their peers. We hypothesized that a larger FN400 will be evident in the group level ERPs for low performers compared to high performers and that there would be a larger LPC evident in the ERPs of high performers compared to low performers.

Methods

Participants. The experiment was conducted as approved by the California State University – San Bernardino Institutional Review Board protocol for research on human subjects**.** Participants were recruited through a combination of methods including advertisements placed around CSUSB or through the school-wide research pool SONA. Participants recruited through advertisements were paid $10 an hour for sessions that lasted approximately. The study consisted of 62 right-handed participants (48 female) whom were students at California State University, San Bernardino and reported being free from neurological and memory problems. Five participants' data were not used due to noncompliance issues and one participant did not have usable data due to technical difficulties during EEG recording. Two participants did not have usable EEG data but were included in behavioral analyses. 56.5% of the participants were self-reported to be Hispanic, 22.6% Caucasian, 11.3% Asian, and 9.7% identified as more than one or another ethnicity. The average age was 23.52 years old (*SD* = 4.82). None of the participants reported any visual, medical, or physical issues that would interfere with the experiment; most spoke English as their first language (N = 47) and the 15 participants who indicated speaking a different language first had been speaking English for an average of 16.73 years (*SD* = 4.74).

Procedure

Participants arrived at the lab and completed informed consent and demographic information forms via voluntary self-report. The paradigm used was a modified item recognition confidence test, which built from similar paradigms successfully used in our lab's prior research (Addante et al., 2012; Addante, Watrous, Yonelinas, Ekstrom, & Ranganath, 2011; Addante, 2015; Addante, Ranganath, Olichney, & Yonelinas, 2012) and described in further detail below (Fig. 1). This paradigm consisted of an encoding phase containing four study sessions, in which participants studied 54 words in each session, and a retrieval phase, containing six test sessions in which the participant's memory was tested for 54 words in each session. They viewed a total of 324 words, 216 of which were presented in the encoding phase and 116 of which were unstudied items (new items).

During encoding, participants were given instructions to make a simple decision about the word presented on the screen. Subjects were asked to either judge if the item was manmade or if the item was alive. The instructions were presented in one of two counterbalanced orders: ABBA or BAAB. The stimuli were presented on a black computer screen in white letters. To begin a trial, a screen with a small white cross at the center was presented for one of three randomly chosen inter-stimuli-interval (ISI) times: 1 second, 2.5 seconds, or 3 seconds. Then, the stimulus word appeared in the middle of the screen with 'YES' presented to the bottom left of the word and 'NO' presented to the bottom right of the word. The participants indicated their answer by pressing buttons corresponding to 'yes' and 'no' with their index and middle fingers, respectively. The response for this screen was self-paced by the participant. After the

participants responded, they viewed a blank black screen at a random duration of 1 second, 2.5 seconds, or 3 seconds. After the blank screen, the small white cross appeared at the center of the screen to begin the next trial. This cycle continued until all 50 words in the all four lists were presented. Between each list, participants were read the instructions for the next task to prevent carryover effects of the preceding encoding task, and to ensure they correctly switched between the animacy and the manmade decision task.

After the encoding phase was complete, the EEG cap was sized while the participant's face was wiped free of skin oil and/or makeup in preparation for attaching ocular electrodes. Five ocular electrodes were applied to the face to record electrooculograms (EOG): two above and below the left eye in line with the pupil to record electrical activity from vertical eye movements, two on each temple to record electrical activity from horizontal eye movements. The EEG cap was placed on the participant's head and prepared for electrical recording. Gel was applied to each cap site and impedances were lowered below 15 KOhms via gentle abrasion to allow the electrodes to obtain a clear electrical signal.

After the EEG cap was in place, the participant began the retrieval phase. The participants were read instructions asking them to judge if the stimulus word presented was old (studied during the encoding phase) or new (not studied before in the encoding phase) (Figure 1). As in the encoding phase, all stimuli words were presented in white font on a black screen. At the beginning of each trial, participants were presented with a word in the middle of the screen, the numbers "1", "2", "3", "4", and "5" evenly spaced beneath the word, the word "New" on the left by the number "1", and the word "Old" on

the right under the number “5”. Participants pressed any number between “1” and “5” to indicate if they confidently believed the word was old (“5”), believe the word was old but was not confident (“4”), did not know if the word was old or new (“3”), believe the word was new but was not confident (“2”), or confidently believed the word was new (“1”) (Addante et al., 2011; 2012a; 2012b); this prompt was subject-paced. Participants were told to choose the response that gave us the most accurate reflection of their memory, and to respond as quickly and accurately as possible. Immediately after the item recognition judgment, participants were asked to answer a source memory confidence test about if the word came from the animacy decision task or the manmade decision task during encoding on a scale of 1 to 5 (Addante et al., 2011; 2012a; 2012b; 2015; Roberts et al., 2018). Participants were instructed to blink only during this blank screen and avoid blinking during the screens with a small cross or stimuli.

Dunning-Kruger In-test Questions. After the source memory test for each 10th word presented during the memory test, the Dunning-Kruger estimate was presented. Participants received instructions asking them to estimate the percentile in which they believed they were performing up to that point in the test compared to other students who would participate in the study. During the test phase, the word “Percentile?” was presented as a prompt for their estimate with the numbers “<60%,”, “60’s”, “70’s”, “80’s”, and “90%+” evenly spaced beneath it. The Dunning-Kruger estimate was participant-paced. During analyses, participants were grouped into quartiles based on their percentile score on the test, allowing us to average each group’s responses and test them against the other group’s average responses to determine significant differences. They were also grouped by errors in percentile estimates; groups of over-estimators,

correct-estimators, and under-estimators (also referred to as estimator groups later) were also made to investigate potential differences in cognitive strategies (see Results).

Dunning-Kruger Post-Test Questions. At the conclusion of the memory retrieval test, participants answered four post-test questions asking them to estimate their objective score on the whole test, their relative percentile on the whole test, how good their memory is in everyday life, and the overall difficulty of the test. (Figure 1). First, they were asked to “Estimate your score on the whole test”. Participants were prompted to respond on a 5-point scale with “1” meaning below 60%, “2” meaning between 60 and 69%, “3” meaning between 70 and 79%, “4” meaning between 80 and 89 percent, and “5” meaning above 90%. The second questions they were asked was the following: “In what percentile did you perform on the whole test?”. The participants were prompted to respond on a 5-point scale with “1” meaning below the 60th percentile, “2” meaning between the 60th and 69th percentile, “3” meaning between the 70th and 79th percentile, “4” meaning between 80th and 89th percentile, and “5” meaning in the 90th percentile or above. The first questions measured perceived objective score on the entire memory test while the second question measured perceived relative score in relation to other students taking the memory test. These post-test prompts allowed us to test for the DKE at a between-subjects level to be sure the effect can be elicited using an episodic memory task.

The two additional post-test questions were: 1) “Rate your memory in everyday life” and 2) “How difficult was this entire test?”. For the first question, participants responded on a 5-point scale with “1” meaning very hard, “2” meaning hard, “3” meaning moderate, “4” meaning easy, and “5” meaning very easy. For the second

prompt, participants responded on a 5-point scale with “1” meaning very bad, “2” meaning bad, “3” meaning moderate, “4” meaning good, and “5” meaning very good (Fig. 1).

Electrophysiological Analyses. Physiological measurements of brain activity were recorded using electroencephalogram (EEG) equipment from Brain Vision, LLC. All EEG data was processed using the ERPLAB toolbox using Matlab (Delorme and Makeig, 2004; Lopez-Calderon & Luck, 2014). The EEG data was first re-referenced to the average of the mastoid electrodes, passed through a high-pass filter at 0.1 hertz (Hz), and then down-sampled to 256 Hz. The EEG data was epoched from 200 milliseconds prior to the onset of the stimulus to 1200 milliseconds after the stimulus was presented, and then categorized based on performance & response accuracy.

Independent components analysis (ICA) was performed using InfoMax techniques in EEGLab (Bell & Sejnowski, 1995) to accomplish artifact correction and then resulting data was individually inspected for artifacts, rejecting trials for eye blinks and other aberrant electrode activity. During ERP averaging, trials exceeding ERP amplitudes of +/- 250 mV were excluded. Using the ERPLAB toolbox (Lopez-Calderon & Luck, 2014), automatic artifact detection for epoched data was also used to identify trials exceeding specified voltages, in a series of sequential steps as noted below. Simple Voltage Threshold identified and removed any voltage below -100ms. The Step-Like Artifact function identified and removed changes of voltage exceeding a specified voltage (100uV in this case) within a specified window (200ms), which are characteristic of blinks and saccades. The Moving Window Peak-to-Peak function is commonly used to identify blinks by finding the difference in amplitude between the most negative and

most positive points in the defined window (200ms) and compared the difference to a specified criterion (100 uV). The Blocking and Flatline function identified periods in which the voltage does not change amplitude within a specified window (848ms). An automatic blink analysis, Blink Rejection (alpha version), used a normalized cross-covariance threshold of 0.7 and a blink width of 400ms to identify and remove blinks (Luck, 2014).

In order to maintain sufficient signal-to-noise ratio (SNR), all comparisons relied upon including only those subjects whom met a criterion of having a minimum number of 12 artifact-free ERP trials per condition being contrasted (Addante, Ranganath, & Yonelinas, 2012; Gruber and Otten, 2010; Kim et al., 2009; Otten et al., 2006; c.f. Luck 2016). ERPs of individual subjects were combined to create a grand average, and mean amplitudes were extracted for statistical analyses. Topographic maps of scalp activity were created to assess the spatial distribution of effects. For ERP figures, a 30hz low pass filter was applied grand averaged ERPs so as to parallel the similar 'smoothing' function that ensues from taking the mean voltage between two latencies during standard statistical analyses (i.e. Addante, 2015).

Results

Behavior. We excluded a total of five participants from behavioral analysis: four were excluded due to non-compliance issues while one was excluded for technical difficulties in recording during the EEG experiment.

Item Memory Performance. Recognition memory response distributions for recognition of old and new items are displayed in Table 1. Item recognition accuracy was calculated as the proportion of hits (M = .81, SD = .11) – proportion of false alarms (M = .24, SD = .14) (i.e. pHit-pFA) (Addante et al., 2011; 2012a, b). Participants performed item recognition at relatively high levels (M = .57, SD = .15) which was greater than chance, $t(55) = 3.59$, $p < .001$. In addition, participants' accuracy for high confidence item recognition trials ('5's') was significantly greater than low confidence item recognition trials ('4's'), $t(55) = 9.04$, $p < .001$.

Source Memory Performance. Source memory response distributions for recognition of old and new items are displayed in Table 2. Source memory accuracy values were collapsed to include high and low source confidence responses which were then divided by the sum of items receiving a correct and incorrect source response to calculate the proportion (Addante et al., 2012a, 2012b; Roberts et al., 2018). Mean accuracy for source memory was .30 (SD = .19) and was reliably greater than chance, $t(55) = 11.78$, $p < .001$.

Dunning-Kruger Performance Judgments

The distribution of responses for each Dunning-Kruger response category for the post-test and in-test Dunning-Kruger responses are shown in Table 3. When plotted against actual performance, results from subjects' reported performance estimates revealed that the canonical Dunning-Kruger Effect was evident in the dataset, thereby replicating the DKE and extending it to our novel episodic memory paradigm (Figure 2). To quantify and analyze this effect, the participants were first split into quartiles based on memory accuracy. The average memory test accuracy, organized by quartile and

each quartile's average post-test Dunning-Kruger response, is listed in Table 3. In order to be able to directly compare participants' post-test relative Dunning-Kruger estimate and their actual percentile, participants' percentile ranking made from their accuracy on the memory test was converted to the 5-point scale of percentile estimates that were used both in-test during the retrieval task and at the end of the experiment.

A difference score for each participant's percentile ranking was calculated by subtracting their post-test relative Dunning-Kruger estimate from their converted percentile ranking mentioned above. The bottom quartile (N=14, M = 2.43, SD = 0.51, $t(26) = 17.69$, $p < .001$), 2nd quartile (N=14, M = 1.79, SD = 0.80, $t(26) = 8.33$, $p < .001$), and 3rd quartile (N=14, M = 1.43, SD = 1.28, $t(26) = 4.16$, $p < .001$) significantly overestimated their percentile ranking while the top quartile significantly underestimated their percentile ranking (N=14, M = -0.79, SD = 0.89, $t(26) = -3.29$, $p = .003$) (Table 4, Figure 3). However, difference scores of the 2nd quartile and the 3rd quartile were not significantly different, $t(26) = 0.88$, $p = .39$, and so the 2nd & 3rd quartile groups were combined in subsequent analyses. The combined 2nd and 3rd quartile group was still significantly overestimated their percentile ranking (M = 1.61, SD = 1.07, $t(54) = 7.98$, $p < .001$). On average, the difference score for the top quartile was significantly different than the score for the bottom quartile ($t(26) = 11.68$, $p < .001$) and the combined 2nd+3rd quartiles ($t(340) = 7.22$, $p < .001$). The difference score for the bottom quartile was also significantly different than the score for the 2nd and 3rd quartiles ($t(38) = 2.93$, $p = .01$).The magnitude of the errors made by each group decreased as percentile increased: the bottom quartile overestimated their percentile by 62.56%, the 2nd quartile overestimated by 37.95%, the 3rd quartile overestimated by 14.56%, and the

top quartile underestimated by 8.30% (Figure 3). This basic finding provides evidence that the DKE was elicited by our memory paradigm- one that to our knowledge has not been shown before, thereby extending the DKE phenomenon to episodic memory.

Response Speed for Dunning-Kruger Judgments.

In order to better investigate differences in cognitive strategies used to make metacognitive judgments in the DK task, participants were next separated into groups based on estimation accuracy instead of percentile ranking; this was based upon their post-test estimates of their relative performance on the memory test (Table 5).[1] Reaction times (RT) had not been able to be measured in previous studies of the Dunning-Kruger Effect because extant studies were limited by including only a single measure of self-estimate of performance at the end of a task, which is not adequate for RT analyses. The current study, however, collected thirty DKE judgements per subject (n=30), which permitted analysis of response times during these phenomena, in order to gain insight into how the different groups might have performed the task differently.

A one-way ANOVA comparing general RTs collapsed across all DKE metacognitive responses (‘1’ through ‘5’) between the three groups of over-estimators (N = 38), under-estimators (N = 10), and correct-estimators (N = 8) revealed no significant differences in response times overall ($F(2,51) = .41$, $p = 0.666$). However, because our hypotheses were specifically interested in possible differences in how the groups were responding in the different response categories for metacognitive

[1] We used the post-test relative Dunning-Kruger estimate to create groups of over-estimators (N = 38), correct-estimators (N = 8), and under-estimators (N = 10), although we also conducted a paired t-test between the average of the in-test Dunning-Kruger responses ($M = 3.14$, $SD = 0.81$) for each person to the post-test relative Dunning-Kruger response ($M = 3.16$, $SD = 0.78$) and found that the two scores did not differ, $t(55) = 1.30$, $p = .20$, justifying the decision to use the post-test relative Dunning-Kruger response to separate our respective groups.

judgements, reaction times were then analyzed to assess if they varied as a function of the specific DKE responses at the two sides of the scale (believing one was performing at the best, and worst, respectively). Our sample size for the under-estimator group (N = 10) constrained to these specific response conditions (‘1’ & ‘5’, respectively) contained only three people (N=3), and though the current paradigm has been previously established as being sensitive to small sample sizes for behavioral and EEG effects of memory (i.e.; N = 3, N = 6; Addante et al, 2012; Addante, 2015) we nevertheless sought to increase the sample size of those who did not exhibit the errors of illusory superiority estimates. Therefore, we collapsed the under-estimator group with the additional group of correct-estimators (N = 2 whom had responses in ‘1’ & ‘5’ categories) to create a larger group for analyses (N=5) [2].

We used a two-factor between-subjects ANOVA to determine if any mean differences existed among the reaction times for Dunning-Kruger groups (over-estimators vs. correct- and under-estimators) while they engaged in judging themselves to be in the highest (response of ‘5’) and lowest (response of ‘1’) estimate groups, respectively. The ANOVA revealed a significant condition x group interaction between Dunning-Kruger groups and responses, $F(1,27) = 8.35$, $p = .008$, which we next explored with planned t-tests. Within-group, the reaction times for over-estimators when rating themselves as performing in the worst (<60th) percentile (*M* = 2204 ms, *SD* = 628 ms) were significantly slower than when the same over-estimators rated themselves in as being in the best 90th percentile (DK response of ‘5’; *M* = 1656 ms, *SD* = 544 ms, N

[2] Note that the under-estimator groups of those who avoided the over-estimating errors of illusory superiority are inherently defined as having limited instances of responding that they believed they were performing in the best percentile ‘5’ responses, and hence naturally reduces the original sample size (N=60) of available subjects with sufficient trials to analyze for these sensitive behavioral analyses.

= 13), $t(21) = 2.24$, $p = .04$. Alternatively, the combined group of correct + under-estimators exhibited RTs with the opposite pattern: showing a slower response time when rating themselves in the 90th percentile or above (*M* = 2457 ms, *SD* = 634 ms, N = 5) and a marginally faster response when rating themselves as performing less than the 60th percentile (*M* = 1604 ms, *SD* = 329 ms, N = 3; $t(6) = -2.12$, $p = .08$) (Figure 4) [3].

Between-groups, over-estimators were significantly faster (*M* = 1656 ms, *SD* = 544 ms) than the collapsed group of accurate- & under-estimators (*M* = 2457 ms, *SD* = 634 ms; $t(18) = -2.56$, $p = .02$) when each were responding that they thought they were doing the best (i.e. in the 90th percentile or above). The converse pattern was evident when rating themselves in the 59th percentile and lower (Table 5), as the over-estimators responded relatively more slowly (*M* = 2178 ms, *SD* = 602 ms, N = 11) than the combined group of correct- and under-estimators (*M* = 1604 ms, *SD* = 330 ms, N = 3), though this general pattern was found not to be significant for the current sample sizes $t(12) = 1.56$, $p = .15$. These results persisted despite the small sample sizes of the groups, and the patterns suggest that future studies using larger groups may find more patterns to be more robust in significance. Nevertheless, the pattern of responding revealed evidence that people who erred to over-estimate their abilities were also responding faster when they believed they were doing the best and slower to say they were the worst, whereas more-accurate under-estimators were slower to say they were the best and relatively quicker to say they were the worst.

[3] When not collapsing across groups, over-estimators' RTs were found to similarly be significantly faster than under-estimators' RTs (M = 2578 ms, SD = 827 ms) while rating themselves in the 90th percentile or above (response of '5', N = 13), t(14) = -2.43, p = .03.

Electrophysiological Results

The EEG data were analyzed in several steps to probe possible differences between metacognitive judgements and cognitive strategies, and as noted in the Methods section, ERP analyses included only subjects whom maintained a minimum number of valid ERP trials for both of the ERP conditions being compared, which resulted in somewhat smaller sample sizes from the original N=60 and is noted in each reported result's degrees of freedom). First, we assessed the data for general ERP differences that could be identified between the tasks of memory and metacognition judgments. To do this, we assessed the ERPs for decisions in all of the Dunning-Kruger judgments collapsed together compared to the ERPs for all item memory judgements collapsed together, to form a standard baseline for comparison since there had not been prior ERP work done in this kind of metacognitive task (Figure 5). This revealed that ERP activity for the metacognitive DKE decisions was significantly greater than those for memory judgements, starting from approximately 300 ms and continuing through 1000 ms at almost every electrode site. These effects were maximal at the central parietal site of Pz through 800ms (300-500 ms: $t(54) = 10.69$, $p < .001$; 400-600 ms: $t(54) = 15.19$, $p < .001$; 600-900 ms: $t(54) = 9.79$, $p < .001$), upon which time the effects became evident as maximal at mid-frontal site Fz from 900-1200 ms ($t(54) = 6.46$, $p < .001$), consistent with prior ERP findings of parietal and anterior P300a/b effects for novelty processing (Curran, 2004; Woodruff et al., 2006; Kishiyama et al., 2004; Knight, 1996; Knight & Scabini, 1998). This basic finding established a foundation

that ERPs during the metacognitive judgments of the DKE were reliably distinct from memory-related activity, which we followed up with further investigation.

Are there differences in how DKE groups were making their memory judgements? We next investigated physiological differences in memory ('old-new' effects of hits - correct rejections) as a function of the different kinds of DK groups (i.e., over-estimators, under-estimators). Memory-related ERP effects ('old-new effects') were analyzed as a function of DK group. When looking early in time (400-600ms) during the latencies characteristic of familiarity-based processing (FN400; Addante et al., 2012a,b; Rugg & Curran, 2007) a 2x2 repeated measures ANOVA with factors of Condition (ERP amplitudes for hits, correct rejections) and Group (over-estimators, under-estimators) at mid-frontal site Fz revealed no significant effects of either factor or interaction. However, at the centro-parietal site Pz there was a significant main effect of Condition ($F(1,46) = 5.63$, $p = .022$) and a reliable condition x group interaction ($F(1,44) = 5.0$, $p = .030$). This was explored with a follow-up between-group t-test, Which found under-estimators had a significantly higher amplitude ($M = 1.39$, $SD = 1.59$) of old-new effects than over-estimators ($M = 0.25$, $SD = 1.38$) that occurred maximally at centro-parietal site Pz but was diffuse across the posterior scalp, $t(44)= 2.24$; $p = .03$. As this difference was evident in the parietal region instead of the expected left frontal region characteristic of the FN400, it may indicate an early activation of recollection activity but does not preclude other possible interpretations of its functional significance.

Later in time, from 600-900ms at left parietal site P3, a 2x2 repeated measures ANOVA with factors of Condition (ERP amplitudes for hits, correct rejections) and Group (over-estimators, under-estimators) from 600-900 at left parietal site P3 revealed

a main effect of Condition ($F(1,46) = 7.36$, $p = .009$), and a Condition x Group interaction ($F(1,44) = 9.91$, $p = .007$). This finding was qualified by a between-group t-test, which revealed that the under-estimator group ($M = 1.96$, $SD = 1.35$) had significantly larger LPC effects (hit – correct rejection amplitudes) than the over-estimator group ($M = 0.30$, $SD = 1.72$), $t(44)= 2.81$; $p = .01$ (Figure 6). This finding suggests that the under-estimator group, which consists of the highest performing individuals, relied on using more recollection than the over-estimator group did in making their memory judgments. Accordingly, since the illusory over-estimators constituted the lowest performing individuals, it is possible that one reason why they performed lower was because of a relative reduction in their reliance upon recollection.

How do metacognitive judgments differ among good and bad, over- and under-estimators? To investigate this core question, we analyzed group-level differences in ERPs between the over- and under-estimators by DKE response (all responses collapsed together). There were significant differences in ERP amplitude between the under-estimators and over-estimators at left-frontal electrode F3 from the early time of 150-250 ms ($M_{Over\text{-}Estimators} = 5.09$, $SD = 3.08$; $M_{Under\text{-}Estimators} = 2.93$, $SD = 2.18$; $t(44) = -2.07$, $p = .04$) and at mid-frontal electrode Fz from 400-600 ms ($M_{Over\text{-}Estimators} = 4.16$, $SD = 5.09$; $M_{Under\text{-}Estimators} = 0.55$, $SD = 4.40$; $t(44) = -2.03$, $p = .048$), such that ERPs for over-estimators/under-performers were far more positive than that of the under-estimators/over-performers (Figure 7). One suggestion from these results is that the frontal effect at 400-600 ms may be characteristic of the FN400 ERP effect related to familiarity-based processing, in that over-estimators may be under-performing because they are relying on the less-specific memory process of familiarity (Yonelinas, 2002;

2010) to guide their self-reflection of past performance in order to make their metacognitive judgments, instead of relying upon the more distinct recollection-related processes (Yonelinas et al., 2010) that evidently appears to be supporting the people who were over-performing/under-estimating (Figure 7).

To help better synthesize and integrate these convergent findings, we plotted these main findings in a summary figure representing the results observed for both metacognition and memory (Figure 8). The group differences during metacognitive judgements about memory performance are consistent with supporting familiarity-based processing in the over-estimators that the under-estimators are evidently not relying upon to make the same judgments. Conversely, the group differences in the parietal 'old-new' effect during memory suggest that people whom later went on to successfully avoid over-estimating their memory performance had made their preceding memory judgments by relying upon the clear contextual details of recollection, whereas those who made the illusory errors of over-estimating their memory performance did not have recollection-related processing supporting those memories and instead relied upon relatively less-accurate familiarity-based heuristics.

Brain-behavior Relationships Between Memory & Metacognition. The preceding results prompted the question of whether there is a systematic relationship evident between memory and metacognition effects at the individual subject level. Across subjects, in groups of both over-estimators and under-estimators, the magnitude of the LPC effect that occurred between 600-900ms at P3 (hits-CR) was found to be significantly correlated with the proportion of Hits made by subjects ($r = .318$, $p = .031$, $N=46$), and also negatively correlated with response speeds for times when high

confidence hits went on to receive correct source memory responses ($r = -.305$, $p = .039$, $N = 46$) (Figure 9). For the under-estimator group, magnitude of the LPC effect was found to be correlated in the negative direction with the average Dunning-Kruger response given in-test by each subject ($r = -.798$, $p = .006$, $N = 10$) but did not exhibit any relationship for the over-estimators ($r = -.014$, $p > .10$, $N = 36$) (Figure 9), indicating that for the higher-performing under-estimators, LPC effects would reliably predict what the ensuing average Dunning-Kruger estimate of that subject would be for their estimated task performance, such that the larger LPC magnitudes predicted relatively lower performance estimates.

Overall, these findings converge to reveal that the larger LPC magnitudes were related to higher proportions of hits on the memory test and with faster response times for recollection-related items of high confidence hits with correct source memory (Addante et al., 2011; Addante et al., 2012a; Roberts et al., 2018). Hence, the LPC was related to recollection, and the more recollection signal (LPC) a subject had predicted the more likely they were to under-estimate their memory performance via their average DK responses in the metacognitive judgments (Figure 9); recollection thus led people to exhibit more humble metacognitive self-awareness. What is special about the under-estimators during their decisions to avoid the pitfalls of illusory superiority? One line of evidence we found was that in the under-estimator group the magnitude of the ERPs for metacognitive judgments from 400-600 ms at mid-frontal site of Fz exhibited a significant positive correlation with the average in-test Dunning-Kruger response given by subjects ($r = .774$, $p = .009$, $N = 10$), though this relationship was not evident for the over-estimators ($r = .025$, $p = .886$, $N = 36$) that exhibited the larger FN400-like effects

(Figure 7, Figure 8), suggesting that the relative lack of familiarity-based processing in the under-estimators appears to be governing them towards reporting a lesser estimation of their performance in the task. That is, without the ambiguity of a familiarity signal it may leave recollection better suited to support these discriminating decisions. Overall, these findings indicate that we can reliably predict how people will perform in either accurately- or inaccurately- self-estimating their abilities by knowing the extent of their recollection-related activity occurring beforehand on a memory task.

DISCUSSION

The current study assessed multiple measures of Dunning-Kruger estimates interspersed throughout an on-going episodic memory test while EEG was recorded. The results from behavioral measures first revealed that the memory paradigm was successful at eliciting the DKE. Participants were separated into performance quartiles and their actual percentile ranking in the group was plotted alongside their estimated percentile ranking (Figure 2). The lowest performing participants in the bottom quartile were found to have drastically overestimated how highly they ranked in their groups while the highest performing participants moderately underestimated their actual ranking. This basic finding was important to identify as a starting point in a novel paradigm for studying the Dunning-Kruger Effect in episodic memory, and its establishment permitted us to continue to explore the data in more specified ways for both behavioral and electrophysiological domains.

Behavioral Findings. The current study's paradigm permitted meaningful collection of reaction times for multiple Dunning-Kruger judgments that could be analyzed at a group level, which prior studies of the DKE have not been able to investigate due to their one-time measures of metacognitive performance estimates at the completion of a study. We found over-estimators were discernably faster than under-estimators in judging themselves to be in the top percentile, but they were slower to judge themselves as being in the bottom percentile; accordingly, under-estimators were relatively slower to report being in the best performance and quicker to claim they were doing poorly. There are several theoretical accounts that could be used to view these results, including elements of cognition, social interactions, and the traditional Dunning-Kruger account of double ignorance (1999).

The first account uses cognition for prototypes to explain the reaction time patterns. Kruger and Dunning's (1999) original results suggest that over-estimators do not understand that they are performing poorly and so they believe they are performing well and placing well within their participant group. This could lead to them having a very positive perception about their ability to perform well on certain types of tasks. Social psychological research on prototypes has shown that answers to questions that are very obviously true (closest to one's prototype) are answered faster (for example, the question, "Is a robin a bird?" will elicit a faster "yes" response than the questions "Is an ostrich a bird?" even though both are true) (Rosch, 1974; Collins & Quillian, 1969). Therefore, if a person's perception of their self (or prototype of themselves) includes that they perform well on tasks, they will be more likely to give a fast response when rating themselves well as opposed to rating themselves poorly. On the other hand, if

they believe they are performing poorly, this perception would oppose the prototype that they have formed, causing them to react slower to rating their performance negatively. The same may be true if under-estimators have formed a perception about themselves that they are only average or even below average: it would then be logical that they would be slower to rate themselves as being best and quicker to believe they are performing poorly. Future research extending this paradigm with inclusion of ‘prototype’ measures of individual subjects would be able to test this hypothesis.

An alternative account of the reaction time results comes from using Kruger and Dunning’s (1999) model of double ignorance of low performers (i.e. 1. they do not know the answer, 2. they do not know they are ignorant of the answer) together with the inability of high performers to estimate their place among their peers due to not realizing the weaknesses of their (non)-peer group. By this account, over-estimators would be fast to report that they are doing well because they believe they are actually doing well, while they are slow to report that they are performing poorly because they do not believe they usually perform poorly or do not want to admit to themselves that they are performing poorly because that would be inconsistent with their metacognitive world-view. Accordingly, the better-performing under-estimators would be competent enough to take pause in responding that they are doing well because they know the ways in which they might have failed (due to ‘competence’), and likewise also would be guided by a humble competence (knowing also what they may not know as well as they could know) to be quicker in believing they were performing poorly. The correlation analysis of RTs (Figure 9) suggest that these judgments may be based upon recollection processing evident from the electrophysiological effects.

Neurophysiological findings. We began exploring the neurophysiology of the DKE by examining brain activity for general differences in processing among the memory and metacognition tasks; that is, assessing the generic extent to which these two judgment types could be established for reflecting different kinds of processing. ERPs between memory trials and self-estimate trials were found to differ reliably beginning from approximately 300 ms into the epoch and continuing throughout the epoch to 1200 ms at almost every electrode site, but being maximal first at posterior parietal sites and then later at mid-frontal regions (Figure 5). This pattern of ERPs are consistent with established properties of P300 ERP effects (or P3a and P3b effects) that are known to have the same distributions of topography and latency of across early/late and posterior/anterior regions, respectively, and which have been well-established as being associated with novelty processing (Dien, Spencer, & Donchin, 2003; Otten & Donchin, 2009; Simons, Graham, Miles, & Chen, 2001). This is consistent with the paradigm in that the DKE judgments were uncommon trials that appeared among the common memory trials in the test, and would have been salient stimuli for eliciting an orienting effect of attention as a novelty item (Kishiyama, Yonelinas, & Lazzara, 2006; Knight, 1996; Knight & Donatella, 1998).

We next explored whether the differential metacognitive judgments were associated with differential ERP patterns. When brain activity of all Dunning-Kruger responses was investigated together, over-estimators (those who performed worst) were found to have a higher mean ERP amplitude than under-estimators at frontal electrode sites during 400-600 ms (Figure 7, 8), consistent with the FN400 effect of familiarity-based processing (Addante et al., 2012a, b; for review see Rugg & Curran,

2007). These early ERP effects suggest that the errors of illusory superiority may be caused by an over-reliance on a generic sense of familiarity similar to as has been found in research on the False Fame Effect (Jacoby et al.,1989a, 1989b, 2004), as opposed to the more specific recollecting of the clear details from their past encounters which would instead provide the contextual cues to guide proper placement of one's perceptual judgments. Under-estimators (those who performed best), on the other hand, exhibited a larger LPC than over-estimators did from 600-900ms during memory judgments (Figures 6, 8), indicating that these humble under-estimators may be making the decisions estimating their performance by reliance upon the clearer details of recollected information, as opposed to the fuzzy sense of familiarity that can come with less accuracy (Yonelinas et al., 2004; 2010). Taken together with the behavioral findings in RTs, it appears that over-estimators were fast to brag, whereas the humbler high-performers were slow to judge themselves as being best- and their caution paid off with better scores. Moreover, ERP data suggested that recollecting the past with clear context and details may be an important part to helping keep us humble, whereas relying upon mere feelings of familiarity (without being sure) may lead us to foolish ends of over-estimating ourselves (though note too that being sure is not sufficient on its own, either; see Introduction review of the anti-correlation of memory and confidence). What may instead delineate on the matter is not confidence, per se, (since one can be highly confident of familiar responses that lack both recollection and accuracy (Addante et al., 2012b, 2012a), but rather the value may be found instead in the construct of recollecting the combination of items-bound-in-context (Diana et al., 2008; Eichenbaum et al., 2007; Addante et al., 2012a), including specific details, and contraindicating information (i.e.

‘non-criterial recollection’, Gallo, Weiss, & Schacter, 2004; Parks, 2007; Yonelinas & Jacoby, 1996) from which to place one’s judgment (and even then, be wary as success is still not guaranteed with mere high confidence (Chua, Hannula, & Ranganath, 2012; Roediger & DeSoto, 2014; Koriat et al., 2008; Nelson & Narens, 1990; Wells, et al., 2002).

In the introduction we postulated that that a framework could be made which provides a memory-based account of Dunning-Kruger Effect for illusory superiority judgements in metacognition, whereby these errors (and successes) can be guided at least in part based upon differences in the cognitive processes of memory familiarity and recollection. In such a model, familiarity would be seen as providing the foundational cognitive processing associated with a heuristic used by people unsure about the details of their past performance on the task and thus guiding them to erroneously over-estimate how well they think they did (over-guessing based upon it feeling familiar but lacking details; Whittlesea, 2002; Whittlesea & Leboe, 2000; Whittlesea & Williams, 2000; Whittlesea, 1993; Whittlesea, Jacoby, & Girard, 1990; Whittlesea & Leboe, 2003). On the other hand, recollection would be seen as the cognitive process supporting people’s abilities to correctly retrieve their memories of the past experiences with richness, detail, and contextual information bound together with the item of the episode (Diana et al., 2008; Eichenbaum et al., 2007). Thus, having the cognitive process of recollection available would guide people to make self-assessments of performance that are more conservatively constrained by the details of the facts of that prior performance, thereby avoiding the some risk of incorrectly assuming an over-performance based merely on it seeming familiar acontextually.

Research on familiarity has identified a contribution of guessing to familiarity judgments that is included in its decision heuristic (Voss et al., 2010; Whittlesea, 2002; Whittlesea & Leboe, 2000; Whittlesea & Williams, 2000; Whittlesea, 1993; Whittlesea, Jacoby, & Girard, 1990; Whittlesea & Leboe, 2003) and that may evidently be leading people here to jump to the wrong conclusions about themselves relative to others, similar to as has been found in the so-called False Fame Effect (Jacoby et al., 1989a, b, 2004). As noted in the Introduction, this concept of familiarity being involved in illusory superiority judgments builds from the 'Reach around knowledge' account (Dunning, 2011), and the data here provide a more concrete substantiation of this within the framework of cognitive processes of episodic memory.

Ideal self-assessments would, theoretically, be those that draw upon correct specific details and context to inform our judgments of self-estimates, which is the characteristic of recollection-based processing (Eichnbaum et al., 2007; Yonelinas et al., 2010; Yonelinas, 2004). This interpretation is, to a certain extent, consistent with prior accounts of differences in people's task competency (Kruger & Dunning 1999; Schlösser et al., 2013; Adams & Adams, 1960; Ehrlinger & Dunning, 2003; Oskamp, 1965; Pennycook et al., 2017; Ryvkin et al., 2012a; C. Sanchez & Dunning, 2017) if differences in illusory superiority judgments are understood as being due to differences in how people encoded the initial mnemonic information either successfully or unsuccessfully into memory. That is, those that did not encode information well (i.e. poor attention, motivation, or distraction; Craik, Luo, & Sakuta, 2010; Craik, Eftekhari, & Binns, 2018; Craik, Naveh-Benjamin, Ishaik, & Anderson, 2000; Fernandes, Moscovitch, Ziegler, & Grady, 2005; Galli, Gebert, & Otten, 2013; Middlebrooks, Kerr, &

Castel, 2017; Weeks & Hasher, 2017) would not be likely to recollect that information later nor then be able to accurately calibrate how well they are actually performing if having to guess with heuristics of general familiarity (Whittlesea, 2002; Whittlesea & Leboe, 2000; Whittlesea & Williams, 2000; Whittlesea, 1993; Whittlesea, Jacoby, & Girard, 1990; Whittlesea & Leboe, 2003). However, as the current study did not investigate activity or measures during the encoding stage of memory, this hypothesis remains to be tested in future research.

Implications

This current experiment provides several novel contributions to the understanding of the Dunning-Kruger Effect. First, to our knowledge, this is the only Dunning-Kruger experiment in which self-estimates on a task relative to a peer group were recorded repeatedly throughout the task, and which include physiological measures. That is, self-estimates in prior studies are usually only acquired once: at the end of the task (Adams & Adams, 1960; Burson, Larrick, & Klayman, 2006; Ehrlinger & Dunning, 2003; Kruger & Dunning, 1999; Oskamp, 1965; Pennycook, Ross, Koehler, & Fugelsang, 2017; Ryvkin, Krajč, & Ortmann, 2012a; C. Sanchez & Dunning, 2017) (although see Simons (2013) for instance of when there was a variation of the task using repeated estimates before the task itself). This innovation to the classic Dunning-Kruger paradigm was critical to collecting both the reaction time measures and neurophysiological measures of ERPs during the metacognitive self-estimates, and offers future researchers a pathway forward to continuing exploring this phenomenon.

Our finding that under-estimators had a larger LPC than over-estimators did gives some insight into the inaccurate estimates of illusory superiority that occur in over-

estimators (relative underperformers). Because the over-estimators had a smaller LPC (in fact, lacked any LPC evidence at all, Figure 7, 8), this suggests that they used less recollection during the episodic memory retrieval task. It is thereby reasonable to deduce that their memories for the episodic performance were diminished accordingly, leading to more metacognitive inaccuracies when trying to the recall episodic events related to their performance based purely upon the availability of familiarity. It is possible that this deficit may also be due to relatively more-impoverished encoding/learning in these subjects leading to relative differences in memory competency among the groups (Dunning, 2011), which can be tested in future studies.

We also found evidence of differences in brain activity between under-estimators and over-estimators when collapsing across all Dunning-Kruger metacognitive responses. Over-estimators had a larger ERP mean amplitude than under-estimators at mid-frontal electrode sites from 400-600ms, which is the characteristic position and latency of the FN400 that has been synonymous with familiarity in many prior studies (Addante et al., 2012; Curran, 2000; Friedman, 2013; Gherman & Philiastides, 2015; Rugg et al., 1998; Rugg & Curran, 2007). In the framework of a memory-related interpretation of these results, one could argue that because we found an FN400 in this condition, over-estimators (under-performers) may have evidently used more familiarity than under-estimators did in making these metacognitive judgments, in lieu of the recollection that under-estimators (over-performers) were evidently relying upon instead. That is, each group was apparently arriving at fundamentally different metacognitive conclusions because they were relying upon, or being influenced by, fundamentally different neurocognitive processes of memory. This was mirrored in the

behavioral data of reaction times, which revealed a crossover interaction pattern of responding: the over-estimators were quick to say they were best but slower to say they were worst, whereas the more 'humble' under-estimators were slower to say they performed best but faster to say they were worst- again suggesting fundamentally different cognitive processing between groups. This matched ERP findings that under-estimators relied upon recollection instead of familiarity, whereas those who assumed they were doing better than they were actually doing relied upon familiarty instead of recollection.

Limitations and Considerations for Future Research

As an ERP study without much directly-related extant research preceding it, the current work maintains inherent limitations of all initial explorations: findings remain to be assessed for generalizability, tested for its boundary conditions, and investigated for replicability across other sample sizes and experimental variables. Some of our findings required relying upon relatively small sample sizes, and though the current paradigm has been previsouly found to be effective in seveal prior studies using even smaller sample sizes of neuropsychological patients (Addante et al., 2012; Addante, 2015), nevertheless, these findings should serve as preliminary discoveries to motivate future work exploring larger group sizes. It should be noted, though, that in attempting to characterize the effects of under-estimators, most people in the study exhibited over-estimating errors, so the majority of our relatievly large sized sample of N = 60 were defined as not being in the under-estimator group that avoided these errors (indeed, this more prevaisve illusory error was the prime focus of the current study and maintained a relatively large ERP sample size). Specified analyses looked at conditions that were

rarely responded to, which further reduced certain sample sizes. Nevertheless, in exploring these effects we maintained rigiorous controls of inclusion criteria for trials to gain effective signal to noise ratio (see Methods), as is attested to by the reliable effects observed in the current work that small sample sizes inherently create a stronger challenge to achieve (which we did).

Additionally, while the electrophysiological results are compelling in suggesting memory effects contributing to the Dunning-Kruger phenomena, other cognitive processes also likely contribute to ERP effects of memory, too, such as implicit fluency and conceptual priming (Voss et al., 2014; Voss & Paller, 2010, 2012; Leynes & Zish; Leynes & Addante, 2016; though see comments in Addante et al., 2012a, 2012b; Mecklinger et al., 2012; Bridger et al., 2012; Bader et al., 2017) and future work would benefit from explorations in those directions as well. It also remains possible, though fully unexplored, that Dunning-Kruger Effects for both behavior & physiology could co-vary on variables of personality such as trait-based features, so futue work integrating combinations of elecotrphysiology, genetics, & personality inventories could prove to be fruitful explorations motivated by the current findings. A final limitation to the current work is that the authors, too, may be inherently subject to the pervasiveness of the DKE and be over-estimating it's value, misinterpreting results, or unaware of counterfactual evidence (indeed, for good measure the authors believe the work to be relatively 'average', and will benfit from the rigors of scientific peer-review/feedback). We hope that the current research can serve in providing value for motivating future research that can investigate these findings in more depth, extend them, and test them against competing hypotheses.

Summary and Conclusions

In conclusion, the current study adds to the literature by a series of small steps: first, it represents the first physiological measures of the DKE, as well as reaction time measures of the phenomenon. Second, the study represents an integrative new paradigm that was developed to permit measuring multiple recurring trials of Dunning-Kruger metacognitive judgments, which others can now use to extend our understanding further. Third, this paradigmatic innovation made possible the ability to capture the DKE in a complex episodic memory task which extends the body of work on the DKE from logic & math problems to episodic memory tasks of item and source memory confidence measures/paradigms.

The current study thus represents an important small step forward in understanding one of the most pervasive observations of human behavior: persistent metacognitive illusions from an error tendency to both over-estimate our abilities as well as to under-estimate our performances. Those who tend to perform best estimate that they under-perform, whereas those who tend to estimate they are the best can realistically perform relatively worse. This pernicious psychological pattern has been observed as far back in history by philosophers such as Socrates and Confucius (Socrates from Apology by Plato, 21d; Confucius, trans. 1938/500), cautioned against by texts of many faiths including Judeo, Christian, & Islamic texts (Isaiah, 5:21; Proverbs 12:15, ; Proverbs 29:11; 1 Corinthians, 3:18; Romans 1:22; Surah Al-A'raf 7:205; Sahih Muslim 2865; Qur'an 25:63; Qur'an 31:18; Qur'an 53:32), noted by laureates such as Shakespeare and scientists such as Darwin, and persists today

throughout the modern age- including in university professors (Cross, 1977). The basic premise of the DKE- that we have inherent illusory errors of our metacognitive self-assessment abilities or performance- is thus seemingly a fundamental force that shapes our socio-psychological universe in much the way that gravity shapes the backdrop of our physical universe: the effect persists through time and affects everyone at some level. Overcoming these psychological errors via metacognitive self-awareness is possible, but like overcoming gravity, it takes a concerted effort requiring careful energy and resources, and then persists only for transient moments in time until returning back to baseline levels. It takes work with persistent self-awareness to avoid the pitfalls of illusory superiority, and surely benefits from practice.

By investigating the underlying neural correlates of this effect, we can begin to categorize exactly why or how such illusory errors of metacognition are occurring. Our finding that over-estimators (under-performers) have reduced recollection-related physiology introduces the possibility of developing countermeasures to improve their memory (such as entrainment devices known to improve memory in our paradigm (Roberts et al., 2018; for Review see Hansylmayer et al., 2019) or non-invasive brain stimulation (such as transcranial direct current stimulation) and investigate if their self-estimates can improve in accuracy with improved memory (Boudewyn, et al., 2018; Cappiello, et al., 2016; Mizrak et al., 2018), or via cognitive interventions of self-awareness training and Cognitive-Behavioral Therapies, for instance. More experimentation is needed, but the present work constitutes an important first step to identifying the cognitive processes involved and preventing the errors that can lead to the leadership & safety hazards of over-confidence.

Together, these innovations revealed convergent insight into why people differ in this phenomenon, evidently based at least in part upon memory-related differences in recollection & familiarity of our past experiences. Group differences in the parietal 'old-new' effect during memory suggest that people whom later went on to successfully avoid over-estimating their memory performance had made their preceding memory judgments by relying upon the clear contextual details of recollection, whereas those who made the illusory errors of over-estimating their memory performance did not have recollection-related processing supporting those memories and instead relied upon relatively less-accurate familiarity-based heuristics. The more recollection signal (LPC) a subject had predicted the more likely they were to under-estimate their memory performance via their average DK responses in the metacognitive judgments (Figure 9); recollection thus led people to exhibit more humble metacognitive self-awareness. The relative lack of familiarity-based processing in the under-estimators appears to be governing them towards reporting a lesser estimation of their performance in the task. That is, without the ambiguity of a familiarity signal it may leave recollection better suited to support these discriminating decisions to better outcomes. Overall, these findings indicate that we can reliably predict how people will perform in either accurately- or inaccurately- self-estimating their abilities by knowing the extent of their recollection-related activity occurring beforehand on a memory task. We hope that this research can serve to inspire new explorations endeavoring to discover the neural correlates of our psychological processes, towards a better understanding of ourselves and the truth of human behavior.

Acknowledgements:
Authors would like to thank Rose DeKock, Raechel Marino, Constance Greenwood, Celene Gonzalez, Yoselin Canizales, Kevin Benitez, Yesensia Casas, Maynori Hinton, Roman Lopez, for helpful assistance in collecting the EEG data, and thank Drs. Donna Garcia & John Clapper for thoughtful feedback on earlier drafts.

This research was supported by the following grants for which the authors acknowledge appreciation for the generous support: National Institute of Health LRP Grant 1 L30 NS112849-01 to RJA from National Institute of Neurological Disorders & Stroke; The Mini-Grant Program from the CSUSB Office of Sponsored Research to RJA; the Faculty Assigned Time Grants from the CSUSB Office of Student Research to RJA; the Faculty Summer Research Fellowships from the CSUSB Office of Academic Research & the Deans Summer Research Fellowship from the College of Social & Behavioral Sciences to RJA; CSUSB Assigned Time for Exceptional Service to Students Award to RJA from Faculty Senate & Provost; the Faculty-Student Research Grants to RJA, LAS, & AM; the CSUSB ASI Student Travel Grants to AM and LAS; the CSUSB Student Success Initiative Innovative Scholars Fund to LAS and AM; and the CSUSB Student Success Initiative Culminating Project Awards to LAS and AM.

Conflict of Interest Statement: Authors report no conflicts of interests.

Author Contributions:
AM collected the data, analyzed data, and co-wrote the manuscript
LS programmed the study, assisted with behavioral data analysis
RJA designed the study, supervised all parts, analyzed data, & wrote the manuscript.

Data Accessibility Statement: Data is accessible upon request.

Abbreviations
DKE: Dunning-Kruger Effect
EEG: Electroencephalogram
ERP: Event-related Potential

REFERENCES

Adams, P. A., & Adams, J. K. (1960). Confidence in the Recognition and Reproduction of Words Difficult to Spell. The American Journal of Psychology, 73(4), 544–552. https://doi.org/10.2307/1419942

Addante, R. J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. NeuroImage, 62(1), 439–450. https://doi.org/10.1016/j.neuroimage.2012.04.031

Addante, R. J., Watrous, A. J., Yonelinas, A. P., Ekstrom, A. D., & Ranganath, C. (2011). Prestimulus theta activity predicts correct source memory retrieval. PNAS Proceedings of the National Academy of Sciences of the United States of America, 108(26), 10702–10707. https://doi.org/10.1073/pnas.1014528108

Addante, R. J. (2015). A critical role of the human hippocampus in an electrophysiological measure of implicit memory. NeuroImage, 109, 515–528. https://doi.org/10.1016/j.neuroimage.2014.12.069

Addante, R. J., de Chastelaine, M., & Rugg, M. D. (2015). Pre-stimulus neural activity predicts successful encoding of inter-item associations. NeuroImage, 105, 21–31. https://doi.org/10.1016/j.neuroimage.2014.10.046

Addante, R. J., Ranganath, C., Olichney, J., & Yonelinas, A. P. (2012). Neurophysiological evidence for a recollection impairment in amnesia patients that leaves familiarity intact. Neuropsychologia, 50(13), 3004–3014. https://doi.org/10.1016/j.neuropsychologia.2012.07.038

Allan, K., Wolf, H. A., Rosenthal, C. R., & Rugg, M. D. (2001). The effect of retrieval cues on post-retrieval monitoring in episodic memory: an electrophysiological study. Brain Res Cogn Brain Res, 12(2), 289-299. doi:10.1016/s0926-6410(01)00061-1

Alicke, M. D., Dunning, D. A., & Krueger, J. (2005). The Self in Social Judgment. Psychology Press.

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, J. I. Krueger, M. D. Alicke, D. A. Dunning, J. I. Krueger (Eds.), The Self in social judgment (pp. 85-106). New York, NY: Psychology Press.

Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. Organizational Behavior and Human Decision Processes, 39(1), 133–144. https://doi.org/10.1016/0749-5978(87)90049-5

Atir, S., Rosenzweig, E., & Dunning, D. (2015). When Knowledge Knows No Bounds: Self-Perceived Expertise Predicts Claims of Impossible Knowledge. Psychological Science, 26(8), 1295–1303. https://doi.org/10.1177/0956797615588195

Bartlett, W. B. (2012). Why the Titanic Sank. Amberley Publishing Limited.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. Psychological Bulletin, 117(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Bohannon, J. N. I., & Symons, V. L. (1992). Flashbulb memories: Confidence, consistency, and quantity. In E. Winograd, U. Neisser, E. Winograd (Ed), & U. Neisser (Ed) (Eds.), Affect and accuracy in recall: Studies of “flashbulb”

memories. (pp. 65–91). New York, NY, US: Cambridge University Press. https://doi.org/10.1017/CBO9780511664069.005
Boudewyn, M., Roberts, B. M., Mizrak, E., Ranganath, C., & Carter, C. S. (2018). Prefrontal Transcranial Direct Current Stimulation (tDCS) Enhances Behavioral and EEG Markers of Proactive Control. Cogn Neurosci. doi:10.1080/17588928.2018.1551869
Bradley, J. V. (1981). Overconfidence in ignorant experts. Bulletin of the Psychonomic Society, 17(2), 82–84. https://doi.org/10.3758/BF03333674
Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. Social Cognition, 4(4), 353-376.
Brown, J. D. (2010). Across the (not so) Great Divide: Cultural similarities in self-evaluative processes. Social and Personality Psychology Compass, 4, 318-330.
Brown, J. D. (2012). Understanding the Better Than Average Effect: Motives (Still) Matter. Personality and Social Psychology Bulletin, 38(2), 209–219. https://doi.org/10.1177/0146167211432763
Brown, R., & Kulik, J. (1977). Flashbulb memories. Cognition, 5(1), 73–99. https://doi.org/10.1016/0010-0277(77)90018-X
Bruett, H., & Leynes, P. A. (2015). Event-related potentials indicate that fluency can be interpreted as familiarity. Neuropsychologia, 78, 41–50. https://doi.org/10.1016/j.neuropsychologia.2015.09.035
Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. Journal of Personality and Social Psychology, 90(1), 60–77. https://doi.org/10.1037/0022-3514.90.1.60
Cappiello, M., Xie, W., David, A., Bikson, M., & Zhang, W. (2016). Transcranial direct current stimulation modulates pattern separation. Neuroreport, 27(11), 826-832. doi:10.1097/wnr.0000000000000621
Carpenter, S. K., Mickes, L., Rahman, S., & Fernandez, C. (2016). The effect of instructor fluency on students' perceptions of instructors, confidence in learning, and actual learning. Journal of Experimental Psychology: Applied, 22(2), 161–172. https://doi.org/10.1037/xap0000077
Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: the role of nonmotivated factors in above-average and comparative-optimism effects. Psychological Bulletin, 130(5), 813-838. http://dx.doi.org/10.1037/0033-2909.130.5.813
Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8(2), 240–247. https://doi.org/10.1016/S0022-5371(69)80069-1
Confucius. (1938). The analects of Confucius. New York: Vintage Books, Random House.
Craik, F. I., Luo, L., & Sakuta, Y. (2010). Effects of aging and divided attention on memory for items and their contexts. Psychol Aging, 25(4), 968-979. doi:10.1037/a0020276
Craik, F. I. M., Eftekhari, E., & Binns, M. A. (2018). Effects of divided attention at encoding and retrieval: Further data. Mem Cognit, 46(8), 1263-1277. doi:10.3758/s13421-018-0835-3

Craik, F. I. M., Naveh-Benjamin, M., Ishaik, G., & Anderson, N. D. (2000). Divided attention during encoding and retrieval: Differential control effects? Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(6), 1744-1749. doi:10.1037//0278-7393.26.6.1744

Critcher, C. R., & Dunning, D. (2009). How chronic self-views influence (and mislead) self-assessments of task performance: Self-views shape bottom-up experiences with the task. Journal of Personality and Social Psychology, 97(6), 931–945. https://doi.org/10.1037/a0017452

Cross, K. P. (1977). Not can, but will college teaching be improved? New Directions for Higher Education, 1977(17), 1–15. https://doi.org/10.1002/he.36919771703

Curran, T. (2000). Brain potentials of recollection and familiarity. Memory & Cognition, 28(6), 923–938. https://doi.org/10.3758/BF03209340

Curran, T., 2004. Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. Neuropsychologia 42, 1088–1106.

Darwin, C. (2009). The Descent of Man, and Selection in Relation to Sex. Place of publication not identified MobileReference.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(4), 730–740. https://doi.org/10.1037/0278-7393.34.4.730

Dien, J., Spencer, K. M., & Donchin, E. (2003). Localization of the event-related potential novelty response as defined by principal components analysis. Cognitive Brain Research, 17(3), 637–650. https://doi.org/10.1016/S0926-6410(03)00188-5

Dulas, M. R., & Duarte, A. (2013). The influence of directed attention at encoding on source memory retrieval in the young and old: an ERP study. Brain Res, 1500, 55-71. doi:10.1016/j.brainres.2013.01.018

Dunning, D. (2011). The Dunning-Kruger effect: On being ignorant of one's own ignorance. In J. M. Olson, M. P. Zanna, J. M. Olson (Ed), & M. P. Zanna (Ed) (Eds.), Advances in experimental social psychology, Vol 44. (pp. 247–296). San Diego, CA, US: Academic Press. https://doi.org/10.1016/B978-0-12-385522-0.00005-6

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. Journal of Personality and Social Psychology, 58(4), 568–581. https://doi.org/10.1037/0022-3514.58.4.568

Düzel, E., Cabeza, R., Picton, T. W., Yonelinas, A. P., Scheich, H., Heinze, H.-J., & Tulving, E. (1999). Task-related and item-related brain processes of memory retrieval. Proceedings of the National Academy of Sciences, 96(4), 1794–1799. https://doi.org/10.1073/pnas.96.4.1794

Dywan, J., & Jacoby, L. (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. Psychology and Aging, 5(3), 379–387. https://doi.org/10.1037/0882-7974.5.3.379

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. Journal of Personality and Social Psychology, 84(1), 5–17. https://doi.org/10.1037/0022-3514.84.1.5

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. Annual Review of Neuroscience, 30, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Epley, N., & Gilovich, T. (2006). The Anchoring-and-Adjustment Heuristic: Why the Adjustments Are Insufficient. Psychological Science, 17(4), 311–318. https://doi.org/10.1111/j.1467-9280.2006.01704.x

Fernandes, M. A., Moscovitch, M., Ziegler, M., & Grady, C. (2005). Brain regions associated with successful and unsuccessful retrieval of verbal episodic memory as revealed by divided attention. Neuropsychologia, 43(8), 1115-1127. doi:10.1016/j.neuropsychologia.2004.11.026

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. Journal of Experimental Psychology: Human Perception and Performance, 3(4), 552–564. https://doi.org/10.1037/0096-1523.3.4.552

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. American Psychologist, 34(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Friedman, D. (2013). The cognitive aging of episodic memory: A view based on the event-related brain potential. Frontiers in Behavioral Neuroscience, 7. https://doi.org/10.3389/fnbeh.2013.00111

Galli, G., Gebert, A. D., & Otten, L. J. (2013). Available processing resources influence encoding-related brain activity before an event. Cortex, 49(8), 2239-2248. doi:10.1016/j.cortex.2012.10.011

Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2004). Reducing false recognition with criterial recollection tests: Distinctiveness heuristic versus criterion shifts. Journal of Memory and Language, 51(3), 473-493. doi:10.1016/j.jml.2004.06.002

Gandolphe, M., & El Haj, M. (2017). Flashbulb memories of the Paris attacks. Scandinavian Journal of Psychology, 58(3), 199–204. https://doi.org/10.1111/sjop.12364

Gherman, S., & Philiastides, M. G. (2015). Neural representations of confidence emerge from the process of decision formation during perceptual choices. NeuroImage, 106, 134–143. https://doi.org/10.1016/j.neuroimage.2014.11.036

Guillory, J. J., & Blankson, A. N. (2017). Using Recently Acquired Knowledge to Self-Assess Understanding in the Classroom. Scholarship of Teaching and Learning in Psychology. https://doi.org/10.1037/stl0000079

Hanslmayr, S., Axmacher, N., & Inman, C. S. (2019). Modulating Human Memory via Entrainment of Brain Oscillations. Trends in Neurosciences, 42(7), 485-499. doi:https://doi.org/10.1016/j.tins.2019.04.004

Harvey, N. (1990). Overconfidence in self-assessment of motor skill performance. In J.-P. Caverni, J.-M. Fabre, M. Gonzalez, J.-P. Caverni (Ed), J.-M. Fabre (Ed), & M. Gonzalez (Ed) (Eds.), Cognitive biases. (pp. 499–509). Oxford, England: North-Holland.

Hayama, H. R., Johnson, J. D., & Rugg, M. D. (2008). The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: specific or non-specific? Neuropsychologia, 46(5), 1211-1223. doi:10.1016/j.neuropsychologia.2007.11.021

Hayama, H. R., & Rugg, M. D. (2009). Right dorsolateral prefrontal cortex is engaged during post-retrieval processing of both episodic and semantic information. Neuropsychologia, 47(12), 2409-2416. doi:10.1016/j.neuropsychologia.2009.04.010

Heaton-Armstrong, A., Shepherd, E., Gudjonsson, G., & Wolchover, D. (2006). Witness testimony: Psychological, investigative and evidential perspectives. (A. Heaton-Armstrong, E. Shepherd, G. Gudjonsson, & D. Wolchover, Eds.). New York, NY, US: Oxford University Press.

Hirst, W., Phelps, E. A., Meksin, R., Vaidya, C. J., Johnson, M. K., Mitchell, K. J., … Olsson, A. (2015). A ten-year follow-up of a study of memory for the attack of September 11, 2001: Flashbulb memories and memories for flashbulb events. Journal of Experimental Psychology: General, 144(3), 604–623. https://doi.org/10.1037/xge0000055

Howell, W. C. (1971). Uncertainty from internal and external sources: A clear case of overconfidence. Journal of Experimental Psychology, 89(2), 240–243. https://doi.org/10.1037/h0031206

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic Cognition: Memory, Perception, and Concept Learning. In G. H. Bower (Ed.), Psychology of Learning and Motivation (Vol. 18, pp. 1–47). Academic Press. https://doi.org/10.1016/S0079-7421(08)60358-8

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. Journal of Personality and Social Psychology, 56(3), 326–338. https://doi.org/10.1037/0022-3514.56.3.326

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. Journal of Experimental Psychology: General, 118(2), 115–125. https://doi.org/10.1037/0096-3445.118.2.115

Jacoby, L. L., Woloshyn, V., & Kelley, C. (2004). Becoming Famous without Being Recognized: Unconscious Influences of Memory Produced by Dividing Attention. In D. A. Balota, E. J. Marsh, D. A. Balota (Ed), & E. J. Marsh (Ed) (Eds.), Cognitive psychology: Key readings. (pp. 322–337). New York, NY, US: Psychology Press.

Khoe, W., Kroll, N. E. A., Yonelinas, A. P., Dobbins, I. G., & Knight, R. T. (2000). The contribution of recollection and familiarity to yes–no and forced-choice recognition tests in healthy subjects and amnesics. Neuropsychologia, 38(10), 1333–1341. https://doi.org/10.1016/S0028-3932(00)00055-5

Kishiyama, M. M., Yonelinas, A. P., & Lazzara, M. M. (2004). The von Restorff Effect in Amnesia: The Contribution of the Hippocampal System to Novelty-Related Memory Enhancements. Journal of Cognitive Neuroscience, 16(1), 15–23. https://doi.org/10.1162/089892904322755511

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. Nature, 383(6597), 256. https://doi.org/10.1038/383256a0
Knight, R. T., & Scabini, D. (1998). Anatomic Bases of Event-Related Potentials and Their Relationship to Novelty Detection in Humans. Journal of Clinical Neurophysiology, 15(1), 3.
Koen, J. D., Borders, A. A., Petzold, M. T., & Yonelinas, A. P. (2017). Visual short-term memory for high resolution associations is impaired in patients with medial temporal lobe damage. Hippocampus, 27(2), 184-193. doi:10.1002/hipo.22682
Kolarik, B. S., Baer, T., Shahlaie, K., Yonelinas, A. P., & Ekstrom, A. D. (2017). Close but no cigar: Spatial precision deficits following medial temporal lobe lesions provide novel insight into theoretical models of navigation and memory. Hippocampus. doi:10.1002/hipo.22801
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. https://doi.org/10.1037/0022-3514.77.6.1121
Kvavilashvili, L., Mirani, J., Schlagman, S., Foley, K., & Kornbrot, D. E. (2009). Consistency of flashbulb memories of September 11 over long delays: Implications for consolidation and wrong time slice hypotheses. Journal of Memory and Language, 61(4), 556–572. https://doi.org/10.1016/j.jml.2009.07.004
Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. Psychological Review, 111(1), 94–110. http://dx.doi.org/10.1037/0033-295X.111.1.94
Leynes, P. A., & Addante, R. J. (2016). Neurophysiological evidence that perceptions of fluency produce mere exposure effects. Cognitive, Affective & Behavioral Neuroscience, 16(4), 754–767. https://doi.org/10.3758/s13415-016-0428-1
Leynes, P. A., & Zish, K. (2012). Event-related potential (ERP) evidence for fluency-based recognition memory. Neuropsychologia, 50(14), 3240–3249. https://doi.org/10.1016/j.neuropsychologia.2012.10.004
Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? Organizational Behavior & Human Performance, 20(2), 159–183. https://doi.org/10.1016/0030-5073(77)90001-0
Loftus, E. F. (1975). Leading questions and the eyewitness report. Cognitive Psychology, 7(4), 560–572. https://doi.org/10.1016/0010-0285(75)90023-7
Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. Journal of Experimental Psychology: Human Learning and Memory, 4(1), 19–31. https://doi.org/10.1037/0278-7393.4.1.19
Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. Psychiatric Annals, 25(12), 720–725. https://doi.org/10.3928/0048-5713-19951201-07
Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. Bulletin of the Psychonomic Society, 5(1), 86–88. https://doi.org/10.3758/BF03336715
Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front Hum Neurosci, 8, 213.

Lord, W. (1986). The night lives on (1st ed.). New York: Morrow.
May, R. S. (1986). Overconfidence as a result of incomplete and wrong knowledge. In R. W. Scholz & R. W. Scholz (Ed) (Eds.), Current issues in West German decision research. (pp. 13–30). New York, NY, US: Peter Lang Publishing.
McCabe, D. P., & Soderstrom, N. C. (2011). Recollection-based prospective metamemory judgments are more accurate than those based on confidence: Judgments of remembering and knowing (JORKS). Journal of Experimental Psychology: General, 140(4), 605–621. https://doi.org/10.1037/a0024014
Mizrak, E., Kim, K., Roberts, B., Ragland, D. J., Carter, C., & Ranganath, C. (2018). Impact of oscillatory tDCS targeting left prefrontal cortex on source memory retrieval. Cogn Neurosci, 9(3-4), 194-207. doi:10.1080/17588928.2018.1512480
Mecklinger, A., von Cramon, D. Y., & Matthes-von Cramon, G. (1998). Event-related potential evidence for a specific recognition memory deficit in adult survivors of cerebral hypoxia. Brain: A Journal of Neurology, 121(10), 1919–1935. https://doi.org/10.1093/brain/121.10.1919
Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017). Selectively Distracted: Divided Attention and Memory for Important Information. Psychol Sci, 28(8), 1103-1115. doi:10.1177/0956797617702502
Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502-517. http://dx.doi.org/10.1037/0033-295X.115.2.502
Morcom, A. M. (2015). Resisting false recognition: An ERP study of lure discrimination. Brain Res, 1624, 336-348. doi:10.1016/j.brainres.2015.07.049
Müller, B. C. N., Tsalas, N. R. H., van Schie, H. T., Meinhardt, J., Proust, J., Sodian, B., & Paulus, M. (2016). Neural correlates of judgments of learning—An ERP study on metacognition. Brain Research, 1652, 170–177. https://doi.org/10.1016/j.brainres.2016.10.005
Nadel, L., & Sinnott-Armstrong, W. P. (2012). Memory and law. (L. Nadel & W. P. Sinnott-Armstrong, Eds.). New York, NY, US: Oxford University Press.
Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd, U. Neisser, E. Winograd (Ed), & U. Neisser (Ed) (Eds.), Affect and accuracy in recall: Studies of “flashbulb” memories. (pp. 9–31). New York, NY, US: Cambridge University Press. https://doi.org/10.1017/CBO9780511664069.003
Nelson, T. O., & Dunlosky, J. (1991). When people’s judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The “delayed-JOL effect.” Psychological Science, 2(4), 267–270. https://doi.org/10.1111/j.1467-9280.1991.tb00147.x
Oskamp, S. (1965). Overconfidence in case-study judgments. Journal of Consulting Psychology, 29(3), 261–265. https://doi.org/10.1037/h0022125
Otten, L. J., & Donchin, E. (2000). Relationship between P300 amplitude and subsequent recall for distinctive events: Dependence on type of distinctiveness attribute. Psychophysiology, 37(5), 644–661.
Parks, C. M. (2007). The Role of Noncriterial Recollection in Estimating Recollection and Familiarity. J Mem Lang, 57(1), 81-100. doi:10.1016/j.jml.2007.03.003
Pena, M. M., Klemfuss, J. Z., Loftus, E. F., & Mindthoff, A. (2017). The effects of exposure to differing amounts of misinformation and source credibility perception

on source monitoring and memory accuracy. Psychology of Consciousness: Theory, Research, and Practice, 4(4), 337–347. https://doi.org/10.1037/cns0000137
Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2017). Dunning–kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. Psychonomic Bulletin & Review. https://doi.org/10.3758/s13423-017-1242-7
Ranganath, C. (2010). A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. Hippocampus, 20(11), 1263–1290. https://doi.org/10.1002/hipo.20852
Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. Neuropsychologia, 42(1), 2–13. https://doi.org/10.1016/j.neuropsychologia.2003.07.006
Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. Journal of Experimental Psychology: Human Perception and Performance, 2(4), 491–502. https://doi.org/10.1037/0096-1523.2.4.491
Roberts, B. M., Clarke, A., Addante, R. J., & Ranganath, C. (2018). Entrainment enhances theta oscillations and improves episodic memory. Cognitive Neuroscience, 9(3–4), 181–193.
Rudy, J. (2013). The Neurobiology of Learning and Memory. 2nd Edition. 2014. Sinauer Associates, Inc.: Sunderland, MA. ISBN: 978-1605352305
Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. Trends in Cognitive Sciences, 11(6), 251–257. https://doi.org/10.1016/j.tics.2007.04.004
Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. Nature, 392(6676), 595–598. https://doi.org/10.1038/33396
Ryvkin, D., Krajč, M., & Ortmann, A. (2012a). Are the unskilled doomed to remain unaware? Journal of Economic Psychology, 33(5), 1012–1031. https://doi.org/10.1016/j.joep.2012.06.003
Ryvkin, D., Krajč, M., & Ortmann, A. (2012b). Are the unskilled doomed to remain unaware? Journal of Economic Psychology, 33(5), 1012–1031. https://doi.org/10.1016/j.joep.2012.06.003
Sanchez, C. A. (2016). Differently confident: Susceptibility to bias in perceptual judgments of size interacts with working memory capacity. Attention, Perception, & Psychophysics, 78(4), 1174–1185. https://doi.org/10.3758/s13414-016-1071-1
Sanchez, C., & Dunning, D. (2017). Overconfidence Among Beginners: Is a Little Learning a Dangerous Thing? Journal of Personality and Social Psychology. https://doi.org/10.1037/pspa0000102
Sanquist, T.F. et al. (1980). Electrocortical signs of levels of processing: perceptual analysis and recognition memory. Psychophysiology 17, 568–576
Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. American Psychologist, 67(8), 603–613. https://doi.org/10.1037/a0029869
Schacter, D. L., & Addis, D. R. (2008). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. In J. Driver, P.

Haggard, T. Shallice, J. Driver (Ed), P. Haggard (Ed), & T. Shallice (Ed) (Eds.), Mental processes in the human brain. (pp. 27–47). New York, NY, US: Oxford University Press.

Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. Neurobiology of Learning and Memory, 117, 14–21. https://doi.org/10.1016/j.nlm.2013.12.008

Schacter, D. L., & Loftus, E. F. (2013). Memory and law: What can cognitive neuroscience contribute? Nature Neuroscience, 16(2), 119–123. https://doi.org/10.1038/nn.3294

Schlösser, T., Dunning, D., Johnson, K. L., & Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the “signal extraction” counterexplanation for the Dunning–Kruger effect in self-evaluation of performance. Journal of Economic Psychology, 39, 85–100. https://doi.org/10.1016/j.joep.2013.07.004

Sedikides, C., Gaertner, L., & Cai, H. (2015). On the panculturality of self-enhancement and self-protection motivation: The case for the universality of self-esteem. *Advances in Motivation Science,* 2, 185–241.

Sen, T., & Boe, W. J. (1991). Confidence and accuracy in judgements using computer displayed information. Behaviour & Information Technology, 10(1), 53–64. https://doi.org/10.1080/01449299108924271

Shakespeare, W. (1998). As you like it. Mineola, N.Y.: Dover Publications.

Shapiro, L. R. (2006). Remembering September 11th: The role of retention interval and rehearsal on flashbulb and event memory. Memory, 14(2), 129–147. https://doi.org/10.1080/09658210544000006

Sieber, J. E. (1979). Confidence estimates on the correctness of constructed and multiple-choice responses. Contemporary Educational Psychology, 4(3), 272–287. https://doi.org/10.1016/0361-476X(79)90047-X

Simons, D. J. (2013). Unskilled and optimistic: Overconfident predictions despite calibrated knowledge of relative skill. Psychonomic Bulletin & Review, 20(3), 601–607. https://doi.org/10.3758/s13423-013-0379-2

Simons, R. F., Graham, F. K., Miles, M. A., & Chen, X. (2001). On the relationship of P3a and the Novelty-P3. Biological Psychology, 56(3), 207–218. https://doi.org/10.1016/S0301-0511(01)00078-3

Skavhaug, I.-M., Wilding, E. L., & Donaldson, D. I. (2010). Judgments of learning do not reduce to memory encoding operations: Event-related potential evidence for distinct metacognitive processes. Brain Research, 1318, 87–95. https://doi.org/10.1016/j.brainres.2009.11.047

Skavhaug, I.-M., Wilding, E. L., & Donaldson, D. I. (2013). Immediate judgments of learning predict subsequent recollection: Evidence from event-related potentials. Journal of Experimental Psychology: Learning, Memory, and Cognition, 39(1), 159–166. https://doi.org/10.1037/a0028885

Smith, M. C., Bibi, U., & Sheard, D. E. (2003). Evidence for the Differential Impact of Time and Emotion on Personal and Event Memories for September 11, 2001. Applied Cognitive Psychology, 17(9), 1047–1055. https://doi.org/10.1002/acp.981

Sommer, W., Heinz, A., Leuthold, H., Matt, J., & Schweinberger, S. R. (1995). Metamemory, distinctiveness, and event-related potentials in recognition memory for faces. Memory & Cognition, 23(1), 1–11. https://doi.org/10.3758/BF03210552

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: A new perspective. Nature Reviews Neuroscience, 8(11), 872–883. https://doi.org/10.1038/nrn2154

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? Acta Psychologica, 47(2), 143–148. https://doi.org/10.1016/0001-6918(81)90005-6

Talarico, J. M., & Rubin, D. C. (2003). Confidence, Not Consistency, Characterizes Flashbulb Memories. Psychological Science, 14(5), 455–461.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Vallone, R. P., Griffin, D. W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. Journal of Personality and Social Psychology, 58(4), 582–592. https://doi.org/10.1037/0022-3514.58.4.582

Walter Lord. (1955). A night to remember. New York: Henry Holt and Company.

Weeks, J. C., & Hasher, L. (2017). Divided attention reduces resistance to distraction at encoding but not retrieval. Psychon Bull Rev, 24(4), 1268-1273. doi:10.3758/s13423-016-1210-7

Whittlesea, B. W. (2002). Two routes to remembering (and another to remembering not). J Exp Psychol Gen, 131(3), 325-348.

Whittlesea, B. W., & Leboe, J. P. (2000). The heuristic basis of remembering and classification: fluency, generation, and resemblance. J Exp Psychol Gen, 129(1), 84-106.

Whittlesea, B. W., & Williams, L. D. (2000). The source of feelings of familiarity: the discrepancy-attribution hypothesis. J Exp Psychol Learn Mem Cogn, 26(3), 547-565.

Whittlesea, B. W. A. (1993). Illusions of familiarity. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(6), 1235-1253. doi:10.1037/0278-7393.19.6.1235

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. Journal of Memory and Language, 29(6), 716-732. doi:http://dx.doi.org/10.1016/0749-596X(90)90045-2

Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). Journal of Memory and Language, 49(1), 62-79. doi:10.1016/s0749-596x(03)00009-3Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. Psychological Review, 114(1), 152–176. https://doi.org/10.1037/0033-295X.114.1.152

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. Psychological Review, 117(4), 1025–1054. https://doi.org/10.1037/a0020874

Wolk, D. A., Schacter, D. L., Berman, A. R., Holcomb, P. J., Daffner, K. R., & Budson, A. E. (2004). An electrophysiological investigation of the relationship between

conceptual fluency and familiarity. Neuroscience Letters, 369(2), 150–155. https://doi.org/10.1016/j.neulet.2004.07.081
Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. Brain Research, 1100(1), 125–135. https://doi.org/10.1016/j.brainres.2006.05.019
Yonelinas, A. P. (1999). Recognition memory ROCs and the dual-process signal-detection model: Comment on Glanzer, Kim, Hilford, and Adams (1999). Journal of Experimental Psychology: Learning, Memory, and Cognition, 25(2), 514–521. https://doi.org/10.1037/0278-7393.25.2.514
Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. Journal of Memory and Language, 46(3), 441–517. https://doi.org/10.1006/jmla.2002.2864
Yonelinas, A. P., Aly, M., Wang, W.-C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. Hippocampus, 20(11), 1178–1194. https://doi.org/10.1002/hipo.20864
Yonelinas, A. P., & Jacoby, L. L. (1996). Noncriterial recollection: familiarity as automatic, irrelevant recollection. Conscious & Cognition, 5(1-2), 131-141.
Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. Psychological Bulletin, 133(5), 800–832. https://doi.org/10.1037/0033-2909.133.5.800
Yonelinas, A. P. (2013). The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. Behavioural Brain Research, 254, 34-44. doi:10.1016/j.bbr.2013.05.030
Yu, S. S., & Rugg, M. D. (2010). Dissociation of the electrophysiological correlates of familiarity strength and item repetition. Brain Research, 1320, 74–84. https://doi.org/10.1016/j.brainres.2009.12.071
Zechmeister, E. B., Rusch, K. M., & Markell, K. A. (1986). Training college students to assess accurately what they know and don’t know. Human Learning: Journal of Practical Research & Applications, 5(1), 3–19.
Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. Nature, 453(7192), 233-235. doi:10.1038/nature06860
Zhang, W., & Luck, S. J. (2015). Opposite effects of capacity load and resolution load on distractor processing. J Exp Psychol Hum Percept Perform, 41(1), 22-27. doi:10.1037/xhp0000013

Tables

Table 1. *Distribution of Responses for Each Item Response as a Proportion of All Memory Responses*

Item Recognition Confidence	1	2	3	4	5
All Old Items	.09	.07	.04	.21	.60
All New Items	.43	.23	.10	.15	.08
Animacy Task	.13	.06	.04	.16	.60
Manmade Task	.08	.04	.03	.14	.71

Table 2. *Distribution of Responses for Each Source Response as a Proportion of All Memory Responses*

Source Recognition Confidence	1	2	3	4	5
All Old Items	.14	.14	.22	.17	.33
All New Items	.05	.08	.70	.09	.08
Animacy Task	.24	.17	.27	.16	.16
Manmade Task	.11	.11	.17	.2	.41

Table 3. *Distribution of Responses for Each Dunning-Kruger Response, as a Proportion of All Dunning-Kruger Responses*

DKE Type	<60%	60-69%	70-79%	80-89%	>90%
In-Test DK Responses	.05	.20	.39	.29	.07
Post-Test DK Responses	.02	.11	.54	.30	.04

Table 4. *Average Recognition Memory Test Accuracy and Average Post-Test and In-Test Dunning-Kruger Relative Response by Quartile*

Quartile	Accuracy	Average Post-Test DK Relative Response	Average In-Test DK Relative Response
Top (N = 14)	.74 (.06)	3.50 (0.65)	3.26 (0.73)
3rd (N = 14)	.62 (.02)	3.29 (0.99)	3.33 (1.01)
2nd (N = 14)	.55 (.04)	2.79 (0.80)	2.79 (0.81)
Bottom (N = 14)	.38 (.08)	3.43 (0.51)	3.17 (0.62)

Note. Standard deviations are in parentheses.

Table 5. *Response Distribution Proportions of Dunning-Kruger Responses and Mean Reaction Times, Standard Deviations, and Sample Size for In-Test Dunning-Kruger Judgments Organized by Estimator Group.*

Group	Dunning-Kruger Judgments	1	2	3	4	5
Over-Estimators	Response Distribution	.05	.19	.39	.28	.09
(n = 36)	Reaction Time	2204	2064	1948	2044	1656
	SD	628	641	644	860	544
	N per Response	10	23	33	27	13
Correct-Estimators	Response Distribution	.09	.28	.33	.25	.05
(n = 8)	Reaction Time	1447	2323	2018	1920	2275
	SD	263	987	890	733	360
	N per Response	2	6	7	5	2
Under-Estimators	Response Distribution	.01	.21	.35	.38	.05
(n = 10)	Reaction Time	1918	2074	2166	1996	2579
	SD	--	1249	543	770	478
	N per Response	1	5	9	9	3
Combined Correct- and Under-Estimators	Response Distribution	0.04	0.24	0.34	0.32	0.05
(n = 18)	Reaction Time	1604	2209	2101	1969	2457
	SD	330	1062	693	729	635
	N per Response	3	11	16	14	5

Note. Means and SD are in milliseconds.

Figure Captions and Figures

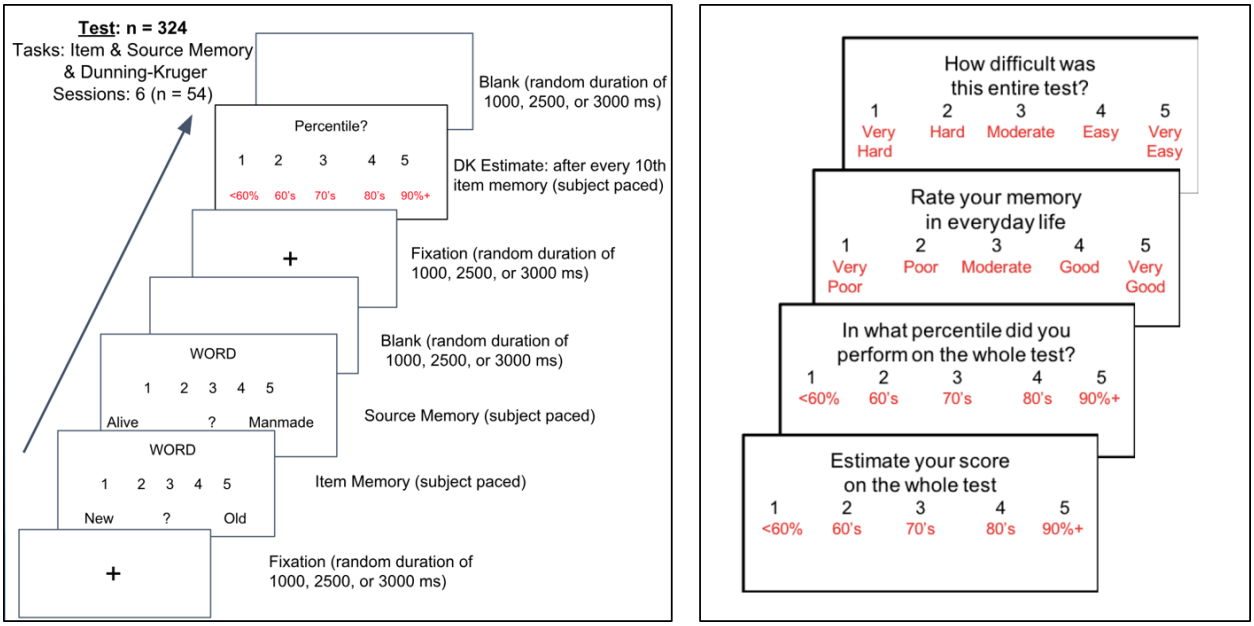

***Figure 1*. Memory Retrieval & Dunning-Kruger Testing Paradigm**. Left: Participants indicated their confidence for item memory and source memory. For every 10th stimuli presented, the participants viewed the Dunning-Kruger Estimate: asking participants to estimate the percentile in which they believe they are performing up to that point on the task in relation to other students. Right: Post-test Dunning-Kruger questions. Participants were asked four questions at the end of the study and responded on a five-point scale with the descriptions seen above.

.

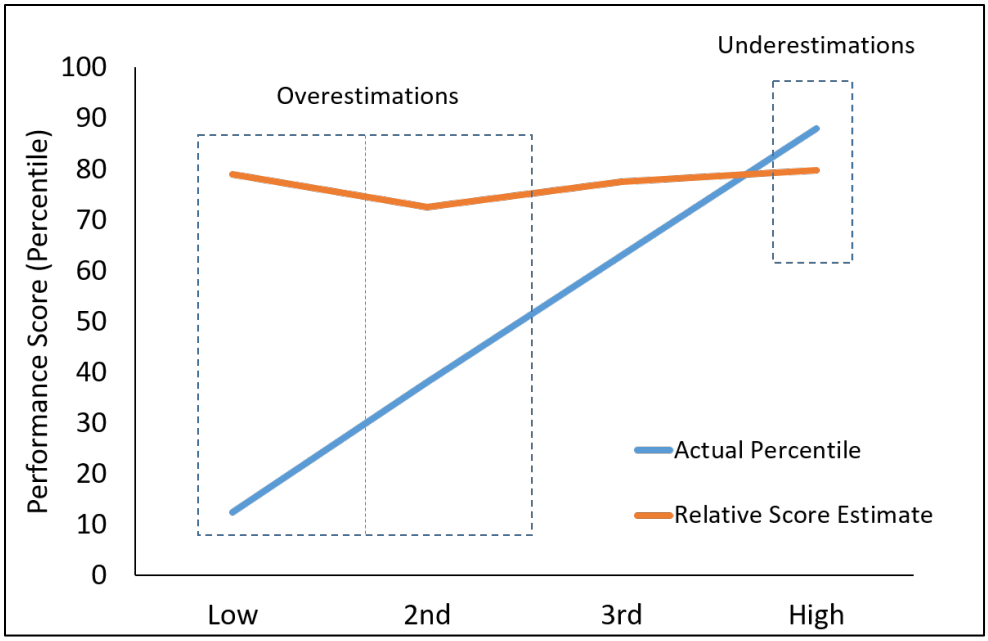

Figure 2. **Actual percentile and estimated percentile by quartile**. Participants were separated by their actual percentile ranking. The low group consists of those in the first quartile (less than or equal to 25%), the second group consists of those in the second quartile (>25% and <=50%), the third group consists of those in the third quartile (>50% and <=75%), and the high group consists of those in fourth quartile (>75%). Participants who performed in the first quartile showed the most overestimation while participants who performed in the fourth quartile showed underestimation of their actual percentile.

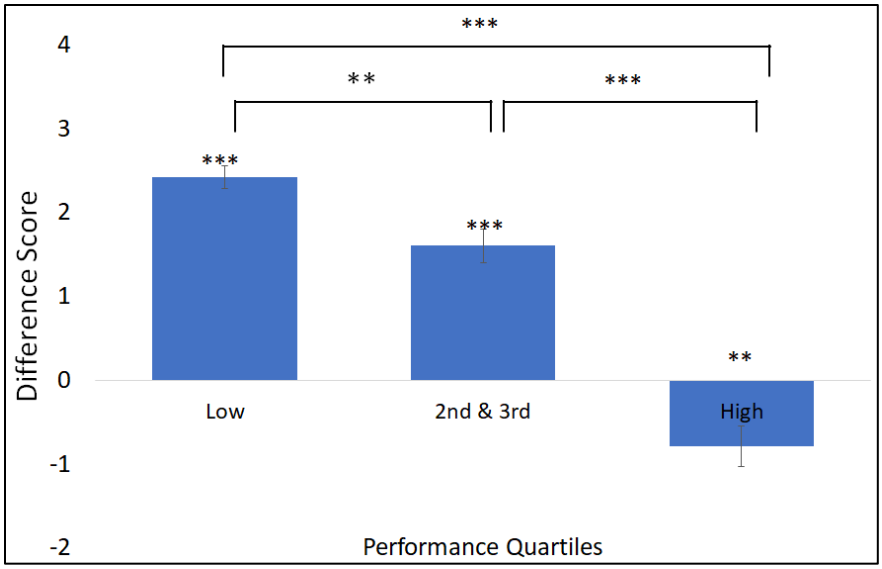

Figure 3. **Average difference score by quartile**. Difference score is calculated by subtracting the converted percentile ranking from the estimated post-test relative score. The bars show the magnitude of overestimation above the x-axis and underestimation below the x-axis for each of the groups. The 2nd and 3rd quartile groups were combined because the groups were not significantly different. Participants in the first quartile and the 2nd & 3rd quartile both overestimated their percentile significantly and were significantly different from each other. Participants in the fourth quartile underestimated their percentile significantly and were significantly different than the low percentile and the 2nd & 3rd percentile. * $p < .05$, ** $p < .01$, *** $p < .001$, † = $p < .10$. Error bars represent the standard error of the mean.

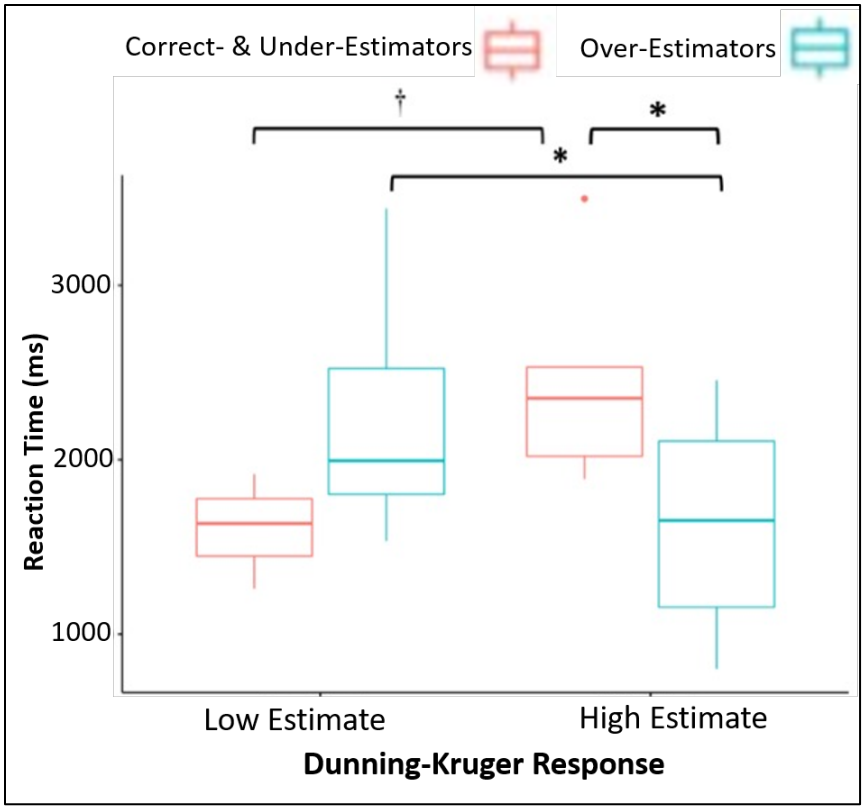

Figure 4. **Mean reaction times of high and low percentile estimation by Dunning-Kruger groups.** Performing in the 59th percentile or below corresponds to response 1 on the task and performing in the 90th percentile or above corresponds to response 5. The reaction times are separated by over-estimators and the combined group of correct- and under-estimators. Mean reaction times are reported in ms. Thick horizontal lines represent means of each graph and the box extent represents inter-quartile range: inside the boxes represent the middle 50% of the data for each variable, respectively, while whiskers represent ranges for the bottom 25% and the top 25% of the data values. * = $p < .05$., † = $p < .10$.

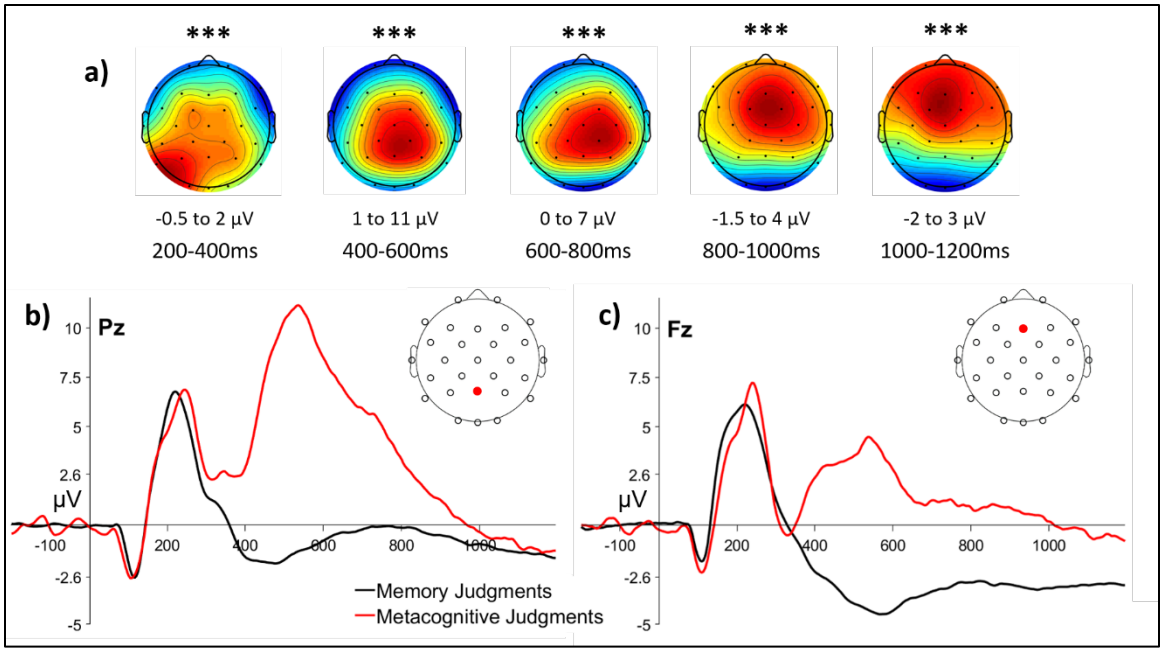

Figure 5. **Comparison of ERPs for memory judgments vs metacognitive judgments estimating performance**. a) Topographic maps of ERPs for all item memory judgments compared to all Dunning-Kruger judgments. Each topographic map is range normalized according to their mean latencies. Warmer colors represent more positive-going voltage differences. b) ERPs for memory and metacognition tasks at central parietal size Pz. c) ERPs for memory and metacognition tasks at mid-frontal site Fz. * $p < .05$, ** $p < .01$, *** $p < .001$.

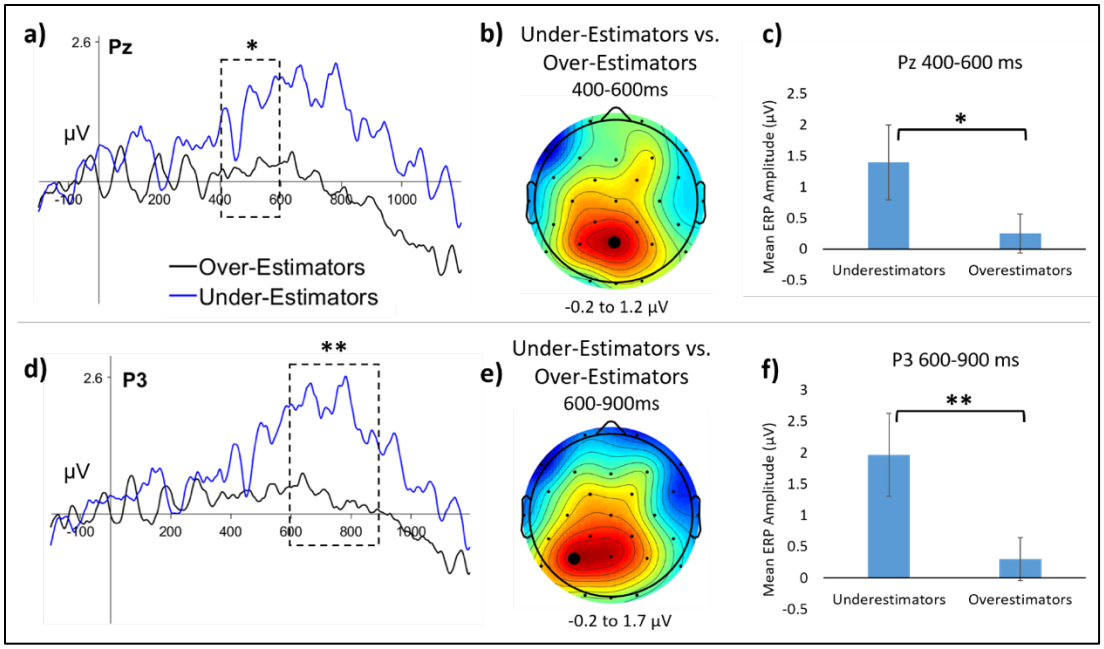

Figure 6. **Difference waves of recognition memory ERP effects for Dunning-Kruger Groups.** Difference waves for memory effects (hits minus correct rejections) for Dunning-Kruger groups of Over- and Under-Estimators at electrode a) Pz and d) P3. The dashed box indicates the latency of the waveform that represents statistically significant effects. Topographic maps show differences in memory effects at a) Pz from 400-600 ms and e) P3 from 600-900 ms. Each topographic map is range normalized according to their mean latencies. Warmer colors represent more positive-going voltage differences. Bar graphs show significant differences in mean ERP amplitude c) Pz from 400-600 ms and f) P3 from 600-900 ms. * $p < .05$, ** $p < .01$, *** $p < .001$.

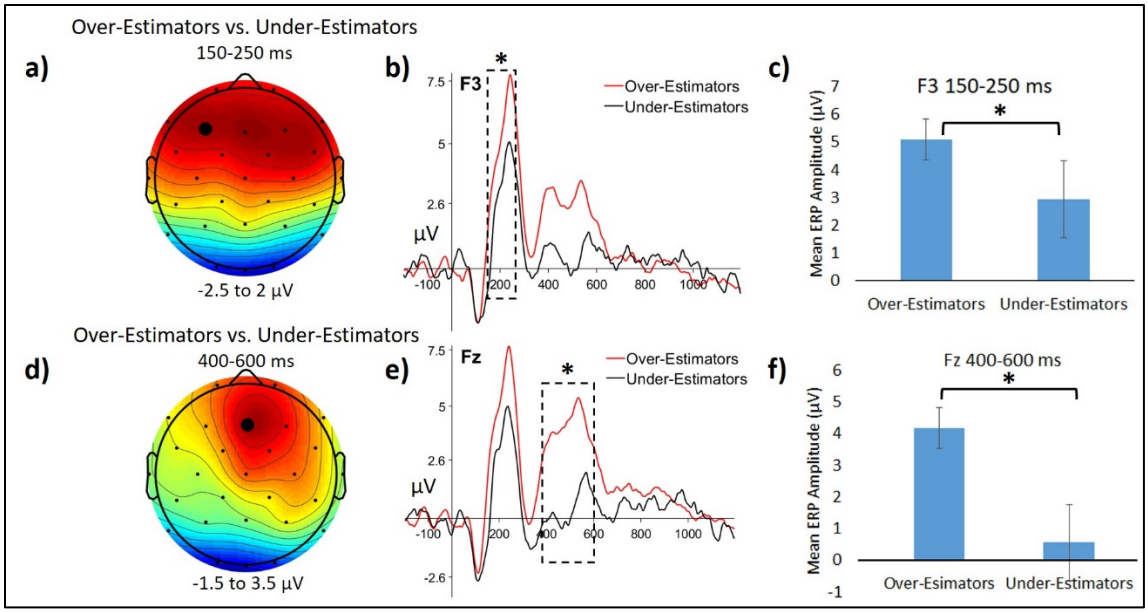

Figure 7. **ERPs of collapsed Dunning-Kruger responses by Dunning-Kruger group.** Topographic maps show ERPs of collapsed Dunning-Kruger responses (Dunning-Kruger judgments 1, 2, 3, 4, and 5 combined) for Over-Estimators compared to Under-Estimators from a) 150-250 ms and b) 400-600 ms and c) for Correct-Estimators compared to Under-Estimators from 800-1200 ms. Each topographic map is range normalized according to their mean latencies. Warmer colors represent more positive-going voltage differences. d), e), and f) show ERPs corresponding to each topographic map to left. g), h), and i) show bar graphs displaying the mean ERP amplitudes corresponding to each ERP and topographic map on the left. The dashed line indicates the time differences that ERPs are significantly different between the groups compared. * $p < .05$, ** $p < .01$, *** $p < .001$.

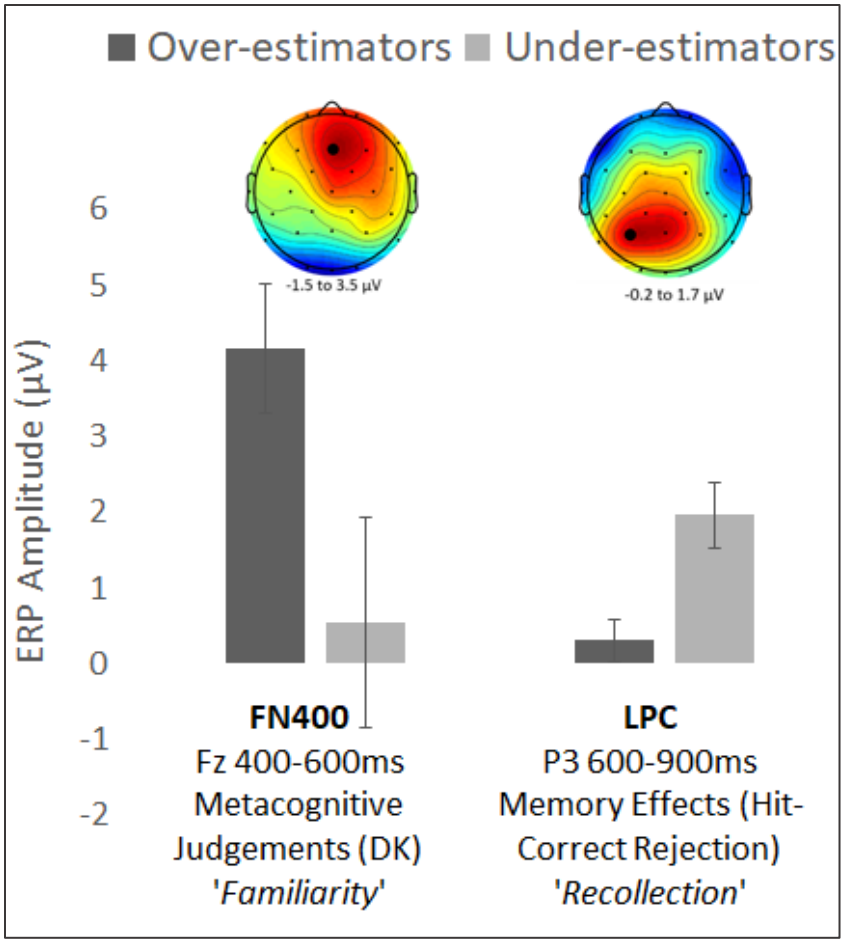

Figure 8. Summary Figure of Dunning-Kruger Effect ERPs. Left: (adapted from Figure 7) Mean amplitude ERP effects for metacognitive judgements (collapsed across all ratings) are shown in the bar graphs for mid-frontal electrode site Fz from 400-600ms and overlaid with the topographic map of the difference wave between the two groups (Over-estimators minus Under-estimators). These group differences during metacognitive judgements about memory performance are consistent with the FN400 effect and suggest familiarity-based processing (Addante et al., 2012a, b; Rugg & Curran, 2007) in the over-estimators that the under-estimators are not relying upon to make the same judgments. Right: (adapted from Figure 6) Mean amplitude ERP 'old-new' effects of item recognition memory (Hits minus Correct Rejections) at left parietal site P3 from 600-900ms, overlaid with the topographic map of the difference between groups (Under-estimators minus Over-estimators). These differences represent the LPC, or parietal 'old-new' effect characteristic of recollection-related memory processing (Addante et al., 2012a, b; Rugg & Curran, 2007), and suggest that people whom later went on to successfully avoid over-estimating their memory performance had made their memory judgments by relying upon the clear contextual details of recollection, whereas those who made the illusory errors of over-estimating their memory performance did not have recollection-related processing supporting those memories and instead relied upon relatively less-accurate familiarity-based heuristics.

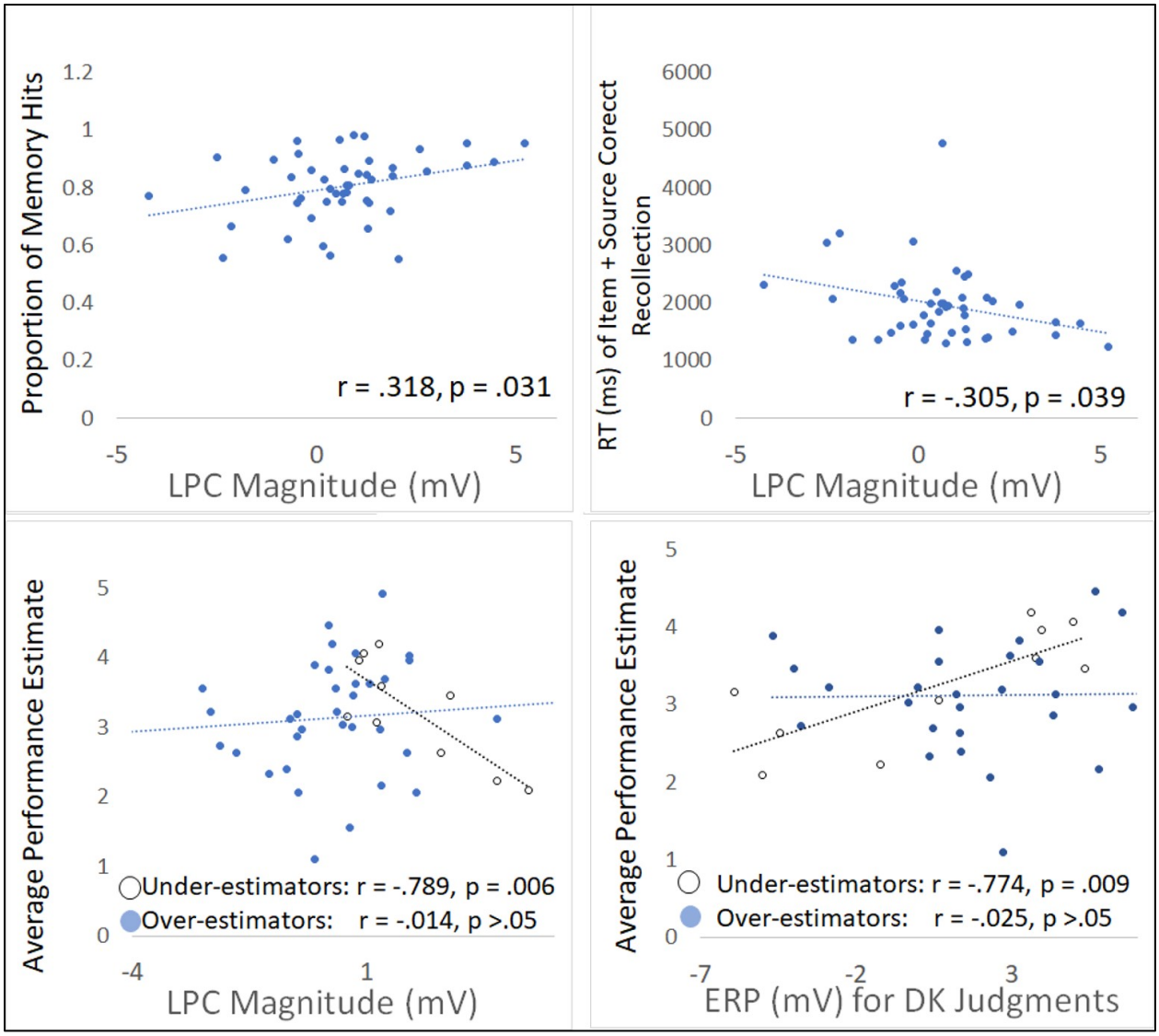

Figure 9. Relationships of Behavioral and Brain Measures for Memory and Metacognition. X-axis represents the magnitude of the LPC effect for both under-estimator and over-estimator groups combined (LPC measured as ERPs for hits minus correct rejections at left parietal site P3 from 600 – 900ms during item recognition memory test (top left, top right, bottom left panels). Bottom right panel x-axis represents the amplitude of mid-frontal ERPs for metacognitive judgments from 400-600ms during the in-test Dunning-Kruger performance estimate task, separated by group. Y-axis represents the proportion of successful item memory hits (judgments of 4 or 5 for 'old' status items during memory retrieval task) (top left); reaction times in milliseconds to recollection-related trials in which subjects got both an high item confidence hit and source memory judgment correct (top right); average in-test performance estimate given by subjects during the metacognitive DK judgments.