This version may not exactly replicate the final chapter
published in the book.
It is not the copy of record!

Metacognition:

Monitoring and Controlling One's Own Knowledge, Reasoning and Decisions

Klaus Fiedler[1], Rakefet Ackerman[2], and Chiara Scarampi[3]

[1]University of Heidelberg (Germany)

[2] Technion—Israel Institute of Technology (Israel)

[3] University College London (United Kingdom)

Running head: Metacognition

Introduction: What is Metacognition?

Setting "Metacognition" Apart from "Cognition"

Metacognition is the "top manager" of cognitive functioning. Memory, for instance, consists of the basic cognitive functions for storing and retrieving information. Metacognitive processes are responsible for regulating these functions: setting goals for learning, examining the quality of memory storage and retrieval, allocating time to memory processes, choosing among strategies for reasoning, making decisions, and achieving goals. Metacognition is not separate from cognition, but integral to all higher-order cognitive inferences, including explicit learning, skill development, recall of personal events, communication, decision making, problem solving, navigation, design, etc. It refers to the superordinate and in a way to the most responsible level of all cognitive functions. It constitutes the quality control of one's own mental functions.

The prefix "meta" in Greek loanwords denotes "something that consciously references or comments upon its own subject" (https://www.dictionary.com/). Thus, metacognition is cognition about one's own cognition. It serves to *monitor* the correctness of our cognitive operations and to correct for incorrect operations in order to *control* for the costs and benefits of our judgments and decisions (Nelson & Narens, 1990). To illustrate, an invoice must be checked (monitoring) and corrected for potential calculation errors (control). Before a written exam can be submitted, all responses must be validated (monitoring) and revised if necessary (control). Purchasing decisions must be confirmed (monitoring) or revised in case of dissatisfying expected results (control).

Monitoring and Control

The output of the monitoring function provides the input to the control function (Nelson & Narens, 1990). Critical assessment of the mental operations used to transform the stimulus information are preconditions for appropriate corrections and for any decisions or actions. Thus, the veracity of verbal communications has to be assessed critically before one can decide whether

to trust, distrust or discard the communication. One monitors the navigation of one's car or boat in order to draw a controlled decision at the next branching. Or, monitoring the money one has spent on prior occasions affords a precondition for the controlled use of the remaining budget.

Metacognition is ubiquitous because virtually all cognitive operations are monitored and controlled, before, during, and after their execution. The execution of an action plan – such as telling a story about what we did last weekend – is not confined to retrieval and speech activities; it also involves monitoring operations such as keeping track of the position reached in the story, checking grammar and pronunciation, assessing the available time left, receiving signals from communication partners, or noting the ease (or difficulty) with which story details come to mind (e.g., "I don't recall the name now, it will probably come up to my mind soon"). As a function of these monitoring results, one can then adjust speed and story detail, correct for mistakes, secure comprehension, and maybe change one's nonverbal behavior in order to appear honest.

Figure 1 provides a schematic overview of generic monitoring and control functions involved in different stages of cognitive processing, from acquisition to retention, retrieval, and inferences leading to judgments and decisions. It is an extended version of a diagram that was originally presented in a seminal article by Nelson and Narens (1990), which was focused on memory processes. As apparent from the direction of arrows, monitoring functions are informed by the contents of the primary cognitive processes, whereas control functions constitute metacognitive influences exerted on the cognitive processes, informed by monitoring results.

Metacognition covers both meta-memory and meta-reasoning (see Ackerman & Thompson, 2015, 2017). That is, monitoring and control functions are not only concerned with memory proper but also with memory-dependent reasoning processes leading to judgments and decision making. Thus, a cognitive-ecological perspective on judgment and decision calls for an extended metacognitive approach, which must not only regulate internal cognitive functions but also check

on the validity and usability of environmental information samples. In this regard, Figure 1

indicates that for judgments and decisions to be unbiased and accurate, the weight given to

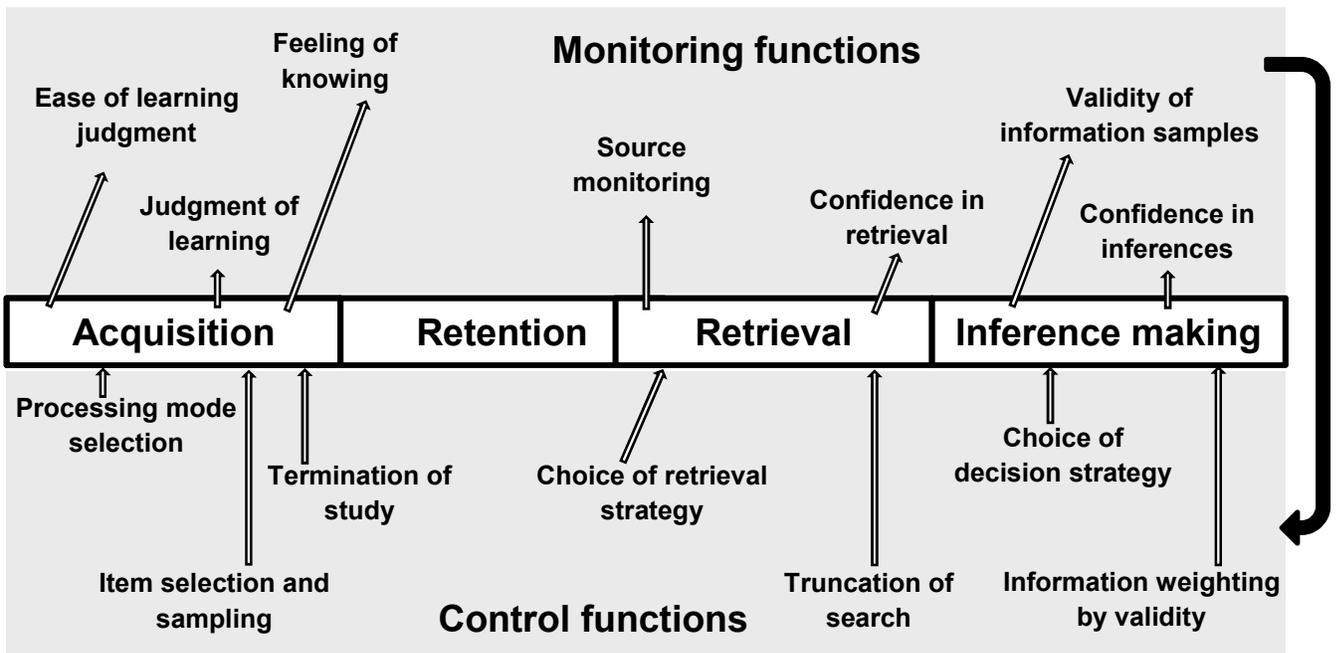sampled information must depend on a critical assessment of its validity and trustworthiness.

Figure 1. Schematic overview of major monitoring and control functions, based on

Nelson and Narens (1990).

Review of Insights Gained from Metacognition Research

In a review of four decades of pertinent research (see Kornell & Bjork, 2007; Son & Sethi,

2010), some milestones can be identified. The general research theme is the interplay of

monitoring and control (Nelson & Narens, 1990), which need not be strictly unidirectional (see

Koriat, Ma'ayan, & Nussinson, 2006). Yet, only when monitoring is reliable can people have a

solid basis for effective control of strategies and allocation of effort and resources.

Metacognitive Regulation of Effort

Imagine Lisa, a student, who is studying a chapter in a text book for an exam. While reading the chapter, Lisa is considering her proficiency and decides whether to repeat looking at a previous paragraph, look for additional information over the Internet, continue to the next paragraph, or stop studying, either because she does not progress adequately today, or because she knows the entire chapter to a satisfactory degree. All these regulatory functions rely on monitoring her knowledge of each paragraph in the chapter. This assessment allows Lisa to identify the weak points and those which she mastered already.

The available evidence on effort regulation was to a large extent collected by simple methodologies involving memorized lists of words or word pairs. Nevertheless, the scientific insights gained from these paradigms are robust and generalizable to many other cognitive tasks (e.g., solving problems, answering knowledge questions, learning from texts, decision making). In a typical paired-associate memory study, people are asked to memorize pairs of related or unrelated words (e.g., KING - CROWN; FLAG – POT) presented one after the other. They are allowed to allocate time to each item freely. Immediately after memorizing each word pair, people assess their chance for success by proving a Judgment of Learning (JOL). For adults, the tasks typically involve memorizing 60 word pairs presented in a random order. After memorizing all of them, there is a recall phase, in which the left words are presented one by one in a new order, and participants are asked to recall the right word that was attached to it in the study phase. Analyses of study time, JOL, and recall success provide evidence about the way people allocate study time across items and in various conditions (e.g., high motivation for success; repeated learning; emotionally loaded vs. neural words; words presented in large or small fonts).

The causal role of JOL for effort regulation was established by Metcalfe and Finn (2008). They asked participants to learn half of the word pairs once and half repeated three times in the

study list. Participants then provided their JOLs and recalled the items. Not surprisingly, JOL and recall were higher for the items learned three times than for those learned only once. In a second block, the items studied once in the first block were now presented three times, and vice versa. All items were thus learned four times altogether and recall of both sets was equivalent. However, JOLs were higher for items learned three times than for those learned only once in the first block, presumably because of the advantage in the initial recall test after the first block. This effect of a previous test on JOL is called *memory for past test*. Most relevant for effort regulation is that when providing JOL for the second block, which differed between the item sets, participants were also asked whether they would like to restudy each item. Although recall performance was equivalent for both item sets, participants chose to restudy items for which JOL was lower – those studied only once in the first block. This finding demonstrates that effort regulation decisions, like decisions to restudy items, depend on JOL rather than on actual memory strength. Similarly, people relied on JOL when betting on success, even when these judgments were misleading (Hembacher & Ghetti, 2017).

Using more complex learning and memory tasks, Thiede, Anderson, and Therriault (2003) found that judgments of comprehension guide decisions to restudy texts. When these JOLs were more reliable, participants were better attuned to their knowledge level and chose to restudy the less well-known texts. This strategy led to higher achievement, demonstrating that effort regulation becomes more effective with more reliable JOLs. For visual perception, subjective confidence guided decisions to get a hint that was helpful for choosing among two options (Desender, Boldt, & Yeung, 2018). Notably, the tight association between monitoring and control was reduced among clinical populations and enhanced among young and healthy people (e.g., Danion, Gokalsing, Robert, Massin-Krauss, & Bacon, 2001; Koren et al., 2004). Thus, a well-functioning monitoring-control link should not be taken for granted.

The next question to ask is when people stop investing efforts. That is, what are the stopping rules that guide effort regulation? A regular finding is that people invest more time in studying the more difficult items (Zacks, 1969). This finding led to the development of *Discrepancy Reduction Models*, which assume that people set a target level according to their motivation in the given scenario. The target acts as a stopping rule: they study each item until monitoring indicates that their knowledge of this item is satisfactory (Nelson & Narens, 1990; see Figure 2). For more difficult items (B in Figure 2) this takes longer than for easier items (A). There are conditions, such as time pressure, under which the stopping criterion gets lower, reflecting a compromise in the target level of knowledge (Thiede & Dunlosky, 1999). High motivation for success, in contrast, leads people to raise their stopping criterion, yielding longer time investment aiming to increase the chances of success (Koriat et al., 2006).

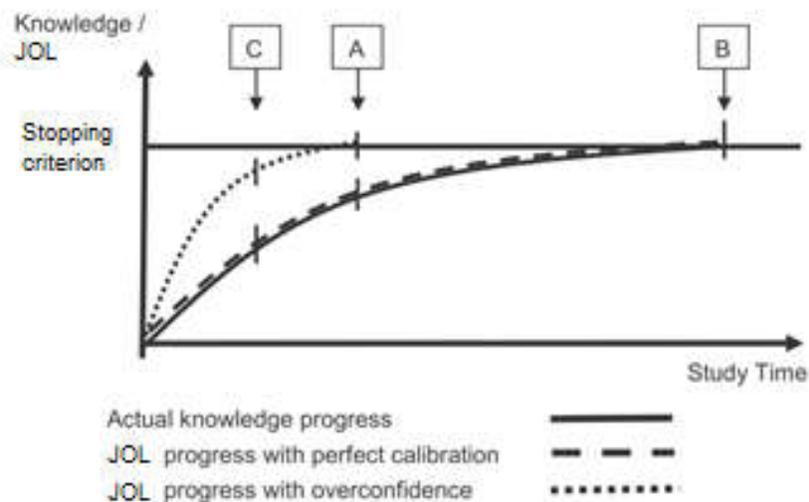*Figure 2.* Illustration of the discrepancy reduction model, based on Ackerman and Goldsmith (2011, Figure 1). It shows the straight criterion and the regulatory role of Judgment of Learning (JOL) in guiding the decision whether to continue or cease learning. A – early termination with overconfidence, B – termination with perfect calibration, C – point of decision to continue learning because the stopping criterion was not reached yet.

As known from real-life scenarios, when the items to be studied are extremely difficult, people may give up early, even when they acknowledge that they do not know them as they would have desired. This strategy is effective since it reduces labor-in-vain: time investment in items that have a low chance of being mastered, even after extensive effort to master them. Moreover, this strategy allows more time to be invested in other items, at intermediate difficulty levels, which have higher chance to be mastered (Son & Sethi, 2010). Indeed, it was shown that people compromise on their target level as more time is invested. They also set a time limit, beyond which they are not willing to invest further time in studying an item (Ackerman, 2014). This time limit is adjusted to be higher when learners have high motivation and to be lower when they learn under time pressure (Undorf & Ackerman, 2017).

One more consideration is order effects. Dunlosky and Ariel (2011) demonstrated that, when presented with several items, people tend to choose to restudy items encountered earlier in their habitual reading order (e.g., from left to right) rather than those appearing later. When study time is too short to master all materials, investing too much in early list parts is counterproductive relative to waiving the most difficult items: The time invested in the difficult items, when they appear early in the list, could be used more effectively for studying easier items appearing later in the list. Beyond this order effect, Dunlosky and Ariel (2011) also found indications for waiving the most difficult items. Thus, these strategies are complementary rather than mutually exclusive.

Generalizing these principles to text learning, Ackerman and Goldsmith (2011) compared learning printed texts to learning the same texts presented on computer screens. In both cases, participants were allowed to write comments and highlight text sections. In the computerized condition, participants believed to learn more quickly than on paper, and thus stopped learning earlier (see Figure 2, point A). In fact, though, rate of learning was equivalent in both media. As a result, performance in tests taken immediately after studying was respectively lower in the

computerized than in the printed-text condition. This apparently reflects the role of

overconfidence in effort regulation—people stop when they think they know the materials

adequately. If they are overconfident, stopping will be premature. Later studies showed that

learning in computerized environments suffers most from limited learning time (for a meta-

analysis, see Delgado, Vargas, Ackerman, & Salmerón, 2018). Similar overconfidence effects

were found with problem-solving tasks of the types students encounter in math, logic, geometry,

and psychometric tests (Ackerman, 2014; Sidi, Shpigelman, Zalmanov, & Ackerman, 2017).

The Heuristic Bases for Metacognitive Judgments

The metacognitive judgments regarding memory, reading comprehension, and solutions to

problems introduced in the preceding section are known to be based on heuristic cues (see

Dunlosky & Tauber, 2014, for a review; Koriat, 1997). Thus, people cannot directly "read" their

knowledge and the quality of their own cognitive processing, but instead, must base their

judgments on cues experienced when they perform the task and immediately after stopping

performing it. One prominent cue is *fluency*—the subjective ease with which a cognitive task is

performed. Fluency is accounted to underlie many metacognitive judgments; it is indeed a rather

valid cue for success. For instance, memorizing the word pair TUBER- AZORES is hard as the

words are rarely encountered and their pairing is rather unusual. When memorizing this word pair

among sixty other pairs, the chances to remember the right word when encountering the left one

remains low despite investing a lot of effort, which means that this item's fluency is low. In

contrast, when a pair consists of familiar words which are often encountered in the same context

(e.g., SOCK – FOOT), cued recall is typically quick and has a high chance of success, and thus

characterized by high fluency. Koriat, Ma'ayan, and Nussinson (2006) suggested that people use

in such contexts a *memorizing effort heuristic*: longer learning times, experienced as lower

fluency, indicate a lower probability of memorizing the item later.

The predictive accuracy of metacognitive judgments depends on the diagnosticity of the utilized cues. A great deal of research focused on conditions under which heuristic cues, like fluency, can be misleading. For instance, people may feel that they found the correct solution for a problem right away and based on fluency be confident they solved it successfully, while in fact they are wrong and investing more effort could increase their chance for success. Thus, identifying factors that induce predictable biases in people's confidence is important because such biases impair effort regulation.

The potentially misleading impact of heuristics suggests that metacognitive judgments are dissociable from actual success of cognitive processes; factors that affect performance do not necessarily affect judgments regarding the same cognitive processes, and vice versa. In particular, dissociation of JOL from actual performance can stem from surface properties of the to-be-learned items affecting perceptual fluency rather than the more relevant cue of processing fluency. Rhodes and Castel (2008) found higher JOLs for words printed in large font than for those printed in smaller fonts, although recall was less affected by font size (see Undorf, Zimdahl, & Bernstein, 2017, for a similar perceptual influence on JOL). Conversely, other variables have more pronounced effects on performance than on JOLs. For instance, rehearsal improves recall and long delays between learning and test cause substantial forgetting, yet JOLs are hardly sensitive to either (Koriat, 1997; Koriat, Bjork, Sheffer, & Bar, 2004). Thus, accuracy of JOLs and other metacognitive judgments depends on the validity of the utilized cues.

An effective and easy-to-adapt solution to several biases of JOLs is to delay the JOL elicitations to a time closer to the test, rather than immediately after learning. The *delayed JOL effect* is robust (see Rhodes & Tauber, 2011, for a meta-analysis). Delayed JOL accuracy reflects access to more diagnostic heuristic cues from long-term memory reflecting better the state of knowledge when taking the test than when provided immediately after learning each item.

In the context of problem solving, Ackerman and Zalmanov (2012) compared performance and confidence in the solutions of multiple-choice and open-ended test format. As expected, they found higher success rates in a multiple-choice test format than in the open-ended test, because of guessing or identifying the correct option when readily available. However, subjective confidence ratings were equivalent in both test formats; they did not reflect this performance difference. Confidence in the same solutions was however sensitive to response time: lower for slow responses than for quick responses. This finding reflects proper utilization of fluency. Similarly, Fernandez-Cruz, Arango-Muñoz, and Volz (2016) found sensitivity to processing fluency for both feeling of error and final confidence in a numerical calculation task. Thompson and colleagues (2013) examined fluency effects on final confidence and on Feeling of Rightness (FOR) — an initial confidence judgment collected immediately after producing the first solution that comes to mind, and before rethinking the solution. They used misleading math problems and considered both processing fluency, based on ease of processing, and perceptual fluency, manipulated by font readability (e.g., hard vs. easy to read fonts). Both FOR and final confidence reflected processing fluency, as both judgments were associated with response times. However, none of the examined judgments reflected perceptual fluency, unlike the aforementioned font-size effects on JOL. This example of a difference between metacognitive judgments of memory processes and of reasoning processes suggests that research should delve into commonalities and differences across tasks (Ackerman & Beller, 2017; Ackerman & Thompson, 2015, for a review).

Convincing evidence for the role of fluency in judgments, as reflected by response time, was provided by Topolinski and Reber (2010). Using three different types of problems, they first presented each problem and then, delayed either for short or longer time, presented a potential answer, which was the target stimulus. Participants had to judge whether the presented answer was the correct solution for the presented problem. For both correct and incorrect candidates,

faster appearing solutions were more frequently judged to be correct than those presented after a delay. Because solution display time was the only difference, the findings indicate that mere delay led to lower endorsement of answers as correct.

Two other heuristic cues were shown to affect feelings-of-knowing regarding answers to knowledge questions. The first cue is the familiarity of the question terms or the knowledge domain (e.g., Reder & Ritter, 1992; Shanks & Serra, 2014). The second cue is accessibility, which reflects the number of associations that come to mind during a retrieval attempt, regardless of whether this information promotes retrieval of correct answers (Koriat, 1993). For example, Koriat and Levy-Sadot (2001) composed general knowledge questions that differed in familiarity of the terms (e.g., the ballets "Swan lake" vs. "The Legend of Joseph") and in accessibility, operationalized as the number of names people can provide for a category (e.g., people tend to know more composers than choreographers). These cues contributed independently to feeling of knowing judgments, which were higher for more familiar objects, especially when items were highly accessible. Accessibility also affected judgments regarding problem solutions (Ackerman & Beller, 2017). Although not necessarily reflected in response time, it is possible that familiarity and accessibility affect fluency by affecting the ease of processing experience.

Focusing on a rarely considered cue, Topolinski, Bakhtiari, and Erle (2016) examined the effects of ease of pronouncing on judgments of solvability—quick assessment whether the problem is solvable—has a solution—or whether it includes contradiction that does not allow solving it at all. Topolinski and colleagues presented participants with solvable anagrams (scrambled words) and unsolvable letter sets that could not be rearranged to form a valid word and manipulated their pronounceability. For instance, for the word EPISODE, they had two anagram options: EDISEPO and IPSDEOE. Easy- and hard-to-pronounce versions also existed for the unsolvable letter sets. As expected, easy to pronounce anagrams were more often rated as

solvable than hard to pronounce anagrams, regardless of whether anagrams were in fact solvable

or not. This finding is particularly interesting because in reality anagrams that are easier to

pronounce are often harder to solve, since people find it harder to rearrange their letters. Thus,

pronounceability may function as a misleading heuristic cue for metacognitive judgments.

Most heuristic cues considered in memory and reasoning research somehow refer to

semantic knowledge activated in verbal tasks. This is the case with relatedness of word pairs,

familiarity of question terms, accessibility of relevant knowledge, and pronounceability, as

reviewed above. Studying heuristic cues that affect perceptual decisions provides opportunities to

consider non-semantic heuristic cues. In a study by Boldt, De Gardelle, and Yeung (2017)

participants judged the average color of an array of eight colored shapes and rated confidence in

their choice. The higher the variability of colors across the eight shapes, the lower was the

participants' confidence in the average color choice, even when equating the actual difficulty.

Thus, people utilize misleading heuristic cues in perceptual tasks as they do in verbal tasks.

When considering the bases for metacognitive judgments, in particular those associated

with fluency, a question in place is whether people base their judgments on the experience of ease

during performing the task (experience-based cues), or on knowledge about cognitive processes,

which is general rather than specific to the current experience with the item at hand (theory-based

cues; Koriat, 1997). For instance, the unjustified effect of font size on JOL mentioned above

could stem from experience of easy learning when the fonts are large relative to an experience of

difficulty when the fonts are small (Undorf & Zimdahl, 2018). The same effect on JOL could also

stem from people's implicit theories of learning, saying that large presentation helps memorizing

while small presentation adds a challenge to the task. Attempts were made to separate the two

information sources. Kelley and Jacoby (1996) aimed to focus on experience-based cues while

controlling for potential theories people might have. They presented participants anagrams

(scrambled words). In the first phase of the experiment, participant studied the solution words to half the anagrams. This prior exposure led to faster solution of those anagrams in a second phase, as the correct solution came to mind more easily. After experiencing such processing ease, participants expected these anagrams to be easier for other people to solve relative to anagrams these participants solved without prior exposure to their answers. This finding demonstrates the intricate role of experience-based cues in metacognitive judgments.

The contribution of experience-based fluency and theory-based beliefs is a source of debate about heuristic cues. Mueller, Tauber, and Dunlosky (2013) found dominance of theory-based beliefs that related word pairs (SOCK - FOOT) were easier to remember than unrelated word pairs (PARROT – GAZ) over effects of experience-based processing fluency on JOLs. Based on Undorf and Erdfelder's (2015) counter-evidence that experience-based fluency is nevertheless an important basis for JOLs, both teams later concluded that theory-based beliefs contribute to JOLs in addition to experience-based fluency (Mueller & Dunlosky, 2017; Undorf & Zimdahl, 2018).

In sum, metacognitive judgments are prone to predictable biases due to utilizing heuristic cues that are generally valid, though misleading under distinct conditions. Understanding factors that people take into account when making metacognitive judgments is essential for any attempt to educate and improve effort and behavior regulation.

Knowing what you Know: Judgment Accuracy

Judgments and decisions are generally accompanied by a subjective feeling of confidence, aimed at assessing the probability of being correct. This metacognitive judgment serves as a guide for current and future behavior, helping people avoid repeating the same mistakes and evaluate whether the available information suffices to make a reliable decision.

Most research on confidence has focused on the relation between confidence judgments and objective performance on a criterion task, with the aim of investigating how well individuals can

monitor their own knowledge. Two main aspects of judgment accuracy can be distinguished, *resolution* (or metacognitive sensitivity) and *calibration* (or metacognitive bias). Resolution refers to distinguishing between correct and incorrect answers (Fleming & Lau, 2014), whereas calibration refers to the extent to which confidence judgments tend to be overconfident (i.e., more optimistic than actual performance) versus underconfident (i.e., less optimistic).

Resolution. Resolution plays an important role in metacognitive control processes and people's behavior (Nelson & Narens, 1990). Imagine a student facing a multiple-choice test in which errors are penalized whereas omissions are not. The test will be solved differently depending on the assessment the student makes of their candidate answers. If an answer is judged as correct it may be worthwhile responding and risking the penalty. In contrast, if an answer is assessed as wrong, the student might decide to withhold the response. The decision to produce or withhold an answer is determined by resolution. Perfect resolution will lead to offer all the candidate responses, which are indeed correct and withhold all incorrect responses. Conversely, poor resolution – at the same level of knowledge – may lead to withhold some of the correct answers and to offer a portion of the incorrect ones, resulting in penalties and lost opportunities for points (Higham & Higham, 2018).

Several indexes of resolution can be computed to assess the accuracy of a judgment once it has been elicited. All measures require the acquisition of an independent performance criterion that quantifies the relationship between accuracy and confidence. In previous research, resolution has been measured using confidence-accuracy correlations within participants (Nelson, 1984). As an alternative, other researchers have suggested signal detection theory (SDT; Green & Swets, 1966; see Figure 4 below), which assesses discrimination between objective states of the world (e.g., distinguishing signal from noise, or the presence or absence of a stimulus). Applied to metacognitive judgments, resolution can be seen as the sensitivity to a signal. More precisely, the

primary cognitive task (e.g., memory, decision making, etc.) is often called Type 1 task, whereas

the task of discriminating of confidence ratings between one's own correct and incorrect

responses in the Type 1 task is called Type 2 task. Advancers of SDT have argued that gamma

correlations can be problematic, as they can be affected by the overall tendency to use higher or

lower confidence ratings (i.e., metacognitive bias; Fleming & Lau, 2014). Nevertheless, gamma

correlations continue to be used in metacognition research. Above-chance confidence-accuracy

correlations were found in a variety of tasks, ranging from perceptual decision making to

challenging problem solving, indicating that people are skilled at identifying whether their

responses are correct or wrong (see Ackerman & Zalmanov, 2012; Koriat, 2018 and references

therein).

Calibration. Another key monitoring accuracy measure in metacognition and self-

regulation is calibration. A simple measure of calibration is the difference between mean

accuracy in task performance and mean confidence. Several studies indicated that people tend to

be overconfident across a variety of conditions (Dunning, Heath, & Suls, 2004). In particular,

Kruger and Dunning (1999) documented a metacognitive bias through which relatively unskilled

individuals not only make erroneous responses, but also overestimate their abilities. That is, a

deficit in knowledge prevents poor performers from distinguishing right from wrong answers and

from realizing how poorly they are performing. However, if trained to become more competent,

their self-assessment becomes more accurate.

Calibration and resolution are independent measures. An individual may have high overall

confidence, but poor resolution and vice versa (Fleming & Lau, 2014). Nevertheless, recent

research has shown that the two are not independent when the probabilistic structure of the

environment is considered (Koriat, 2018). Across a series of experiments using two-alternative

forced choice items from different domains (e.g., perceptual decision making, general

knowledge, memory, and predictions about others' judgments, beliefs, and attitudes), Koriat (2018) found that resolution is strictly dependent on accuracy of Type 1 task performance and that positive correlations between confidence and accuracy observed across many studies are confined to items for which accuracy is better than chance. Furthermore, calibration depended on task difficulty: items with accuracy smaller than 50% led to a strong overconfidence bias, whereas items for which accuracy was better than chance were associated with almost perfect calibration. These results support the proposition that for difficult items that are likely to elicit erroneous responses, individuals are largely unaware of making a mistake. Consistent with this account, the overconfidence bias decreases markedly when the selective reliance on difficult items is avoided through representative sampling (Gigerenzer, Hoffrage, & Kleinbölting, 1991).

Another key element of metacognitive judgments is the time of elicitation. Judgments can be prospective (i.e., occurring before performing a task), or retrospective (i.e., occurring after task completion). For example, a student may reflect on their current knowledge to predict their success on an upcoming test (prospective judgment) and, judge afterwards how well they did, trying to estimate their grade (retrospective judgment). Few behavioral studies have pitted prospective against retrospective judgments for the same task. Siedlecka, Paulewicz and Wierzchoń (2016) compared prospective and retrospective confidence judgments. Participants rated whether presented words were the solution to anagram tasks. Participants also rated their certainty, either before or after seeing the suggested solution. The authors found that post-decision confidence ratings were more accurate than ratings made prospectively. Resolution and calibration were also found to be higher in retrospective than in prospective judgments by Fleming, Massoni, Gajdos, and Vergnaud (2016), using a perceptual decision task. Retrospective confidence ratings were provided on every trial, whereas prospective judgments were only provided prior to every fifth trial. The authors found dissociable influences on prospective and

retrospective judgments. Whereas retrospective judgments were strongly influenced by current-trial fluency, and accuracy and confidence in the immediately preceding decision, prospective judgments were influenced by previous confidence over a longer time window. Furthermore, individual overconfidence was stable across prospective and retrospective judgments, suggesting that overconfidence represents a stable personality trait (Ais, Zylberberg, Barttfeld, & Sigman, 2016; Jackson & Kleitman, 2014).

As many reasoning and problem-solving tasks go on over an extended period of time, the assessment of performance and success probability must be updated repeatedly (Ackerman, 2014). Intermediate confidence is an internal estimate of the adequacy of possible responses considered before arriving at a final solution (see Ackerman & Thompson, 2017). To study this process, Ackerman (2014) asked participants to rate their intermediate confidence every few seconds until they provided a solution, after which they rated their final confidence. The first intervening judgment turned out to be a good predictor of the amount of time participants spent solving the problems. Confidence tended to increase over time. However, whereas at the beginning, participants tended to provide answers when confidence was high, over time they became more willing to provide answers at a lower level of confidence. Final low-confidence responses could be as low as 20%, even when there was an option to give up, by answering "I don't know".

The study of confidence judgments has been extended in the last decades to collective decision making. In numerous perceptual as well as cognitive decisions, interacting individuals can make more accurate decisions by discussing one's own perceptual experiences with others and integrating different opinions, achieving a reliable collective benefit even in the absence of objective feedback (Bahrami et al., 2010). That is, the accuracy achieved by sharing and combining subjective information via social interaction can exceed the accuracy of each

individual opinion, even that of the best individual in the group. This phenomenon is known as "two-heads-better-than-one" effect (Koriat, 2012) or "wisdom of the crowd" (Surowiecki, 2004). Koriat (2012) presented participants with two-alternative forced inference tasks and showed that members of a dyad can take advantage of the wisdom of the group by using a simple heuristic: choosing the response expressed with the highest level of confidence. These findings have relevant implications for collective and democratic decisions and actions.

Neuroscience of Metacognition

In recent years, the study of metacognition was enriched by growing evidence from neuroscience concerning the underlying neurocognitive architecture. Specific neural substrates (especially in frontolateral, frontomedial, and parietal regions; see Figure 3) are involved in metacognition (e.g., Fleming, Huijgen, & Dolan, 2012; Fleming, Ryu, Golfinos, & Blackmon, 2014; Fleming, Weil, Nagy, Dolan, & Rees, 2010). However, the neural bases of human metacognition remain controversial. Metacognition operates on a variety of first-order processes, ranging from memory to perception, problem solving, etc. The diversity of the tasks to be monitored and controlled complicates the study of its neural signature, as it can be difficult to differentiate between the neural activations attributable to the metacognitive monitoring and control processes and the neural signature of the first-order cognitive/emotional processes (Metcalfe & Schwartz, 2016).

Existing attempts to isolate the metacognitive monitoring and control processes from first-order processes, testify to the uniqueness of metacognitive processes. Initial evidence was obtained from neuropsychological cases. For instance, Shimamura and Squire (1986) suggested that frontal lobe (behind forehead) impairments in patients with Korsakoff's syndrome – a chronic memory disorder characterized by severe anterograde amnesia – can impact metacognitive judgments independently of cognitive performance per se. A common finding

suggests that neural signals involved in error monitoring originate in posterior medial frontal

cortex (pMFC; Dehaene, Posner, & Tucker, 1994).

Since the introduction of these seminal studies, further research in such domains as

memory, perception, and decision making, has identified neural correlates of metacognitive

judgments and further dissociated cognitive from metacognitive processes. Fleming et al. (2010)

had participants performing a perceptual decision-making task and providing ratings of

confidence after each decision. The authors found considerable variation between participants in

metacognitive accuracy. Using MRI, this variation in confidence accuracy was found to be

correlated with grey matter volume in the right rostrolateral areas of the prefrontal cortex (PFC).

Furthermore, greater accuracy in metacognitive judgments was associated with increased white-

matter microstructure connected with this area of the PFC. These results point to neural bases of

metacognition that differ from those supporting primary perception. Similarly, in a study by Do

Lam and colleagues (2012), participants who had first learned the pairwise associations between

faces and names were then presented again with each face and asked to provide judgments of

learning (JOLs) regarding chance to recall the associated name. A neurological dissociation was

found between the processes of memory retrieval, which were located in the hippocampal region

(i.e., medial temporal lobes), and those underlying JOLs, which were located in medial PFC,

orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC).

Anatomical, functional, and neuropsychological studies have confirmed the consistent

involvement of a frontoparietal network in metacognition (Vaccaro & Fleming, 2018).

Activations were located in posterior medial PFC, ventromedial PFC and bilateral anterior PFC/

dorsolateral PFC. Other researchers observed activations in bilateral insula and dorsal precuneus

(Vaccaro & Fleming, 2018). These results suggest that the parietal cortex, particularly precuneus,

and insula represent key nodes supporting metacognition, together with the PFC.
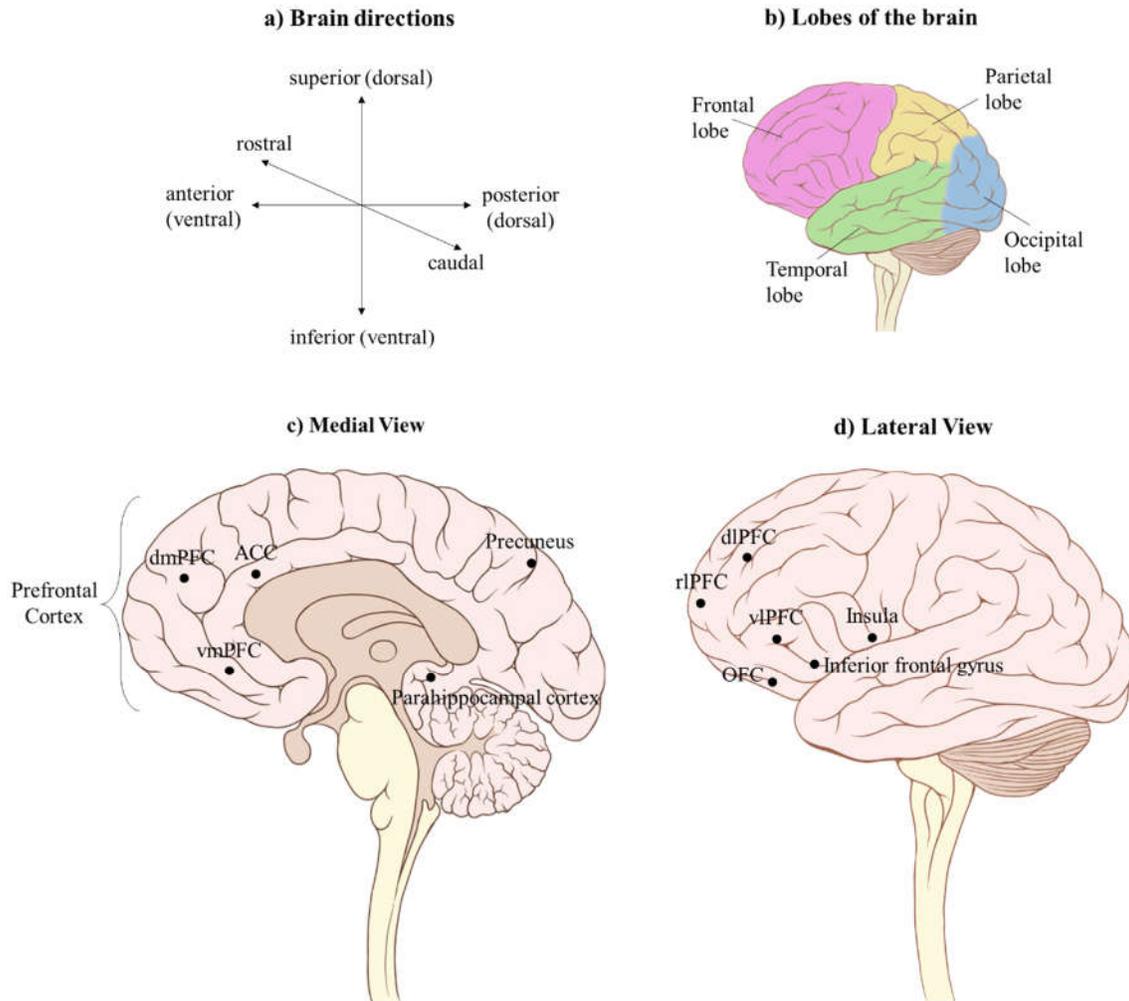
Figure 3. Gross neuroanatomy. *A*, Relative position and direction of brain structures. *B*, The four brain lobes from a lateral view. *C* and *D*, Approximate locations of the broadest subdivisions of the PFC and other areas linked to metacognition. Illustrations adapted from Patrick J. Lynch, medical illustrator, C. Carl Jaffe, MD, cardiologist, under the Creative Commons Attribution 2.5 License, 2006 (CC-BY-2.5). Retrieved from https://commons.wikimedia.org/wiki/File:Brain_human_lateral_view.svg and https://commons. wikimedia.org/wiki/File:Brain_human_sagittal_section.svg. *Abbreviations*: dmPFC, dorsomedial prefrontal cortex; vmPFC, ventromedial prefrontal cortex; ACC, anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; rlPFC, rostrolateral prefrontal; vlPFC, ventrolateral prefrontal cortex; OFC, orbitofrontal cortex.

Existing research supports the existence of neural dissociations between prospective and retrospective metacognitive judgments (Chua, Schacter, & Sperling, 2009; Fleming & Dolan, 2012). For example, in a study on patients with lateral frontal lesions, Pannu, Kaszniak, and Rapcsak (2005) found impaired retrospective confidence judgments, but preserved judgments of future task performance. Conversely, Schnyer and colleagues (2004) found an association between damage to the right ventromedial PFC and a decrease in accuracy for metacognitive judgments about future recall (feeling of knowing), but not for accuracy of retrospective confidence judgments. Further evidence comes from functional MRI studies, which have shown that prospective metacognition activates medial aspects of the PFC, while retrospective metacognitive accuracy is correlated with lateral PFC activity (Fleming & Dolan, 2012). When separating metamemory judgments by temporal focus in their meta-analysis, Vaccaro and Fleming (2018) found that retrospective judgments were associated with activity in bilateral parahippocampal cortex and left inferior frontal gyrus, whereas prospective judgments activated posterior medial PFC, left dorsolateral PFC, and right insula. Nevertheless, neuroimaging evidence directly comparing between different judgement types is scarce. In one of the few studies directly comparing neural activation related to prospective feeling of knowing and retrospective confidence judgment, Chua and colleagues (2009) found an association between prospective judgements and activation in medial parietal and medial temporal lobe, whereas retrospective judgements were associated with inferior prefrontal activity. However, common activations associated with both prospective and retrospective judgments were also observed in regions of medial and lateral PFC, and mid-posterior areas of cingulate cortex. These results suggest that neural activations related to different judgment type may differ in degree rather than in kind (Vaccaro & Fleming, 2018).

Another relevant question tackled in neuroscience is whether metacognition relies on a common, domain-general resource or on domain-specific components that are peculiar to the respective first-order tasks. Recent neuroimaging studies yielded pertinent evidence for both domain-general and domain-specific neural markers (see Rouault, Mcwilliams, Allen, & Fleming, 2018, for a review). A frontoparietal network contributes to metacognitive judgments across a range of different domains. Still, neuroimaging evidence for direct comparisons is scarce. In a recent meta-analysis, Vaccaro and Fleming (2018) observed common regions in separate investigations of memory and decision-making task, which included: insula, lateral PFC, and posterior medial PFC. As suggested by Morales et al. (2018), this result may indicate that judgments in both memory and decision making are driven by common inputs. The meta-analysis also pointed to further regions that are activated by specific tasks. More precisely, meta-memory engaged left dorsolateral PFC and clusters in bilateral parahippocampal cortex, whereas right anterior dorsolateral PFC was involved in decision making (Vaccaro & Fleming, 2018).

In summary, the neural underpinnings of even the most straightforward metacognitive judgments are complicated. Although metacognition can be dissociated from task performance, most studies have revealed activations in multiple brain areas, and differences have emerged between prospective and retrospective judgments. Convergent evidence indicates that the function of the rostral and dorsal areas of the lateral PFC is important for the accuracy of retrospective judgments of performance. In contrast, prospective judgments of performance seem to depend on medial PFC. Recent studies have resulted in a rather nuanced picture, suggesting the co-existence in the brain of both domain-specific and domain-general signals.

Metacognitive Perspectives on Applied Rationality

The research reviewed so far has proven to be fruitful and thought-provoking, suggesting metacognitive explanations bases of adaptive behavior. We have seen that metacognitive deficits

can lead to irrationality and inefficiency. In particular, we have reviewed memorable evidence on the illusion of knowledge, which consists of the gross overestimation of one's chance of success, typically brought about by deceptive feelings of fluency or flow (Fiedler, 2013). Overconfidence, in particular, can be a major source of bias and a dangerous obstacle in decision making under risk and under uncertainty (Glaser & Weber, 2007; Kruger & Dunning, 1999).

The metacognitive perspective is of particular importance for applied research on rational thinking, adaptive regulation, medical diagnosis and treatment, democratic decision making, lie detection, debunking of fake news, argumentation, trust, (im)moral action, and procedural justice in courtrooms, selection committees, or executive decisions. Checking and optimizing the quality of higher-order cognitive operations – the very domain of metacognition – is crucial for rational and responsible behavior. We illustrate this point in the remainder of this section.

Legal Judgments and Decisions

A classical domain of metacognitive research in legal psychology is eyewitness identification performance. Because everybody expects eyewitnesses to identify the perpetrator in a lineup and because the persons in the presented lineup are much more vivid than the original persons in a past episode, the enhanced readiness to make a positive recognition decision produces many correct identifications (when the identified suspect is indeed the perpetrator) but also many incorrect identifications (when the identified suspect is not the perpetrator).

As illustrated in Figure 4, a liberal identification criterion (rather left position of C) produces, say, 90% correct identifications but roughly 40% incorrect identifications. A high false-alarm rate can be conceived as a case of overconfidence; C is apparently too weak a criterion to discriminate guilty and innocent persons, yielding an intolerably high rate of wrong convictions. Consistent with this account, a majority of exoneration cases after the introduction of DNA proofs turned out to be innocent victims of incorrect eyewitness identification.
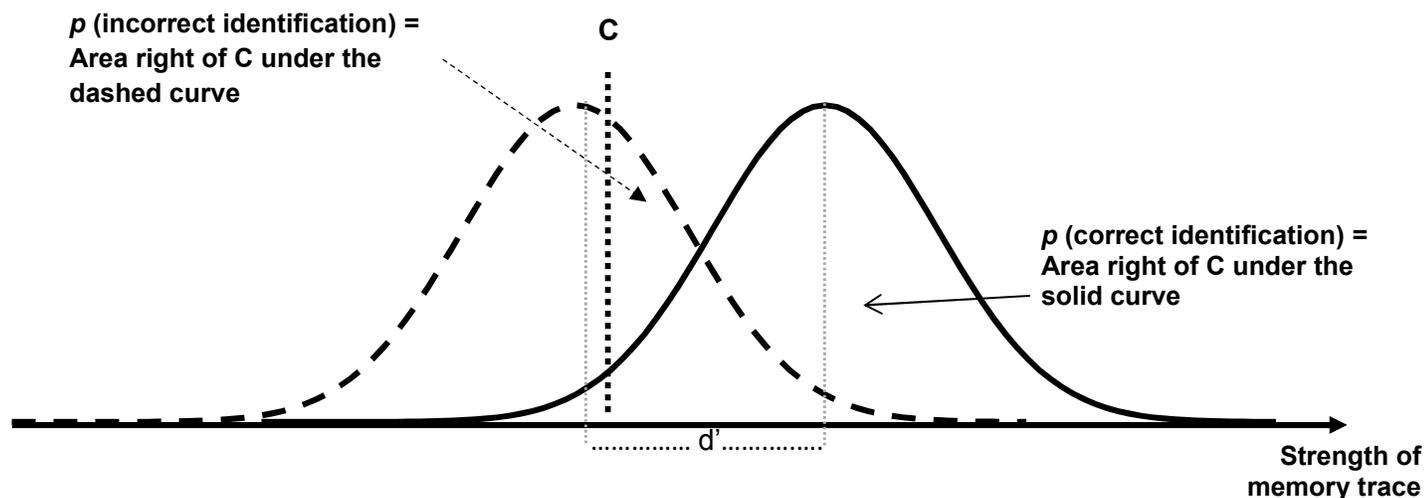
Figure 4. Signal detection analysis of eyewitness-identification performance: The solid (dashed)

curve represents the distribution of memory strength when the suspect in a lineup is (is not) the

real perpetrator. Discriminability is the average horizontal difference d' between curves. An

identification decision is made when memory strength exceeds the criterion C. The areas right of

C under the solid (dashed) curve are the probabilities of correct (incorrect) identification.

The distinction between prospective and retrospective confidence judgments is also relevant

to eyewitness testimony (Nguyen, Abed, & Pezdek, 2018). Witnesses are often asked to rate

shortly after witnessing a crime their ability to recognize the perpetrator in the future (prospective

confidence). Subsequently, when asked to identify someone from a lineup, eyewitnesses are

asked how confident they are that they identified the correct person as perpetrator (retrospective

confidence). Nguyen, Abed, and Pezdek (2018) found that postdictive confidence was a better

indicator of identification accuracy than predictive confidence, both for faces of the same race as

the witness and for cross-race faces. Consistent with the lab findings reviewed above, this

suggests that eyewitness confidence should be collected at the time of identification rather than

earlier on the crime scene.

Metacognitive Myopia as a Major Impediment of Rationality

Less optimistic insights were obtained in other areas of metacognition research. Rational judgments and decisions about economic, political, legal, and health-related issues rely heavily on the critical assessment of both the logical correctness of mental operations and the validity of the underlying evidence. A conspicuous deficit in this sorely needed function of critical assessment has been termed *metacognitive myopia* (Fiedler, 2000, 2008, 2012). As the term "myopia" (short-sightedness) suggests, experimentally demonstrated violations of rational norms typically do not reflect insufficient attention or insensitivity to the stimulus data. On the contrary, people are quite sensitive to the data given; they are in a way too sensitive, taking the data for granted and failing to discriminate between valid and invalid information. For example, when judging the success of different stocks on the stock-market, participants were quite sensitive to the frequency with which various stocks were reported in TV programs among the daily winners. However, they failed to take into account that the daily winning outcomes of some stocks had been reported in more than one TV program (Unkelbach, Fiedler, & Freytag, 2007). Although they fully understand that two TV programs on the same day provide the same stock-market news, participants do not exhibit much success in taking the redundancy into account. Even when they are explicitly reminded of the redundancy and instructed not to be misled by such repetitions, they cannot avoid their misleading influence. This failure to overcome a known pitfall is a metacognitive flaw.

Analogous findings were observed across many experimental tasks. Fully irrelevant numerical anchors influence quantitative judgments (Wilson, Houston, Etling, & Brekke, 1996). Samples that dramatically over-represent the base-rate of rare events (e.g., samples in which the prevalence of HIV is 50% rather than 0.1% as in reality) are used to estimate associated risks (Fiedler, Hütter, Schott, & Kutzner, 2018). Correctly denied questions referring to objects or

behaviors not included in a film nevertheless increased the probability that the non-existing

objects were later recalled erroneously (Fiedler, Armbruster, Nickel, Walther, & Asbeck, 1996).

In a perseverance paradigm, explicit debriefing about an experimental lie did not erase the

implications and psychological consequences of the lie (L. Ross, Lepper, & Hubbard, 1975).

Common to all these findings is that participants, who fully understand that invalid stimulus

information should be discarded, are nevertheless influenced by that invalid information.

The conspicuous naivety with which information is used and retained uncritically,

regardless of its invalidity, is reminiscent of Hannah Arendt's (1963) admonition that compliance

and uncritical conformity are the origin of severe harm and violations of legal norms of

humanity. But although the super-ego residing in metacognition's pre-frontal brain area is

ethically obliged to engage in critical test and reconfirmation, its role in higher-order cognition is

often impoverished. Meta-analyses of modern research on debunking (Chan, Jones, Hall

Jamieson, & Albarracín, 2017), for instance, testify to the inability of scientific or political

debriefing to erase fake-news or obvious myths. Thus, even when the public are fully debriefed

that Iraq did not possess any atomic bombs when the US invaded, that the evidence on global

warming is uncontestable, or that polygraph lie detection is not supported by reliable studies,

people change their erroneous beliefs only slightly and continue to hold the discredited wrong

beliefs to a considerable extent.

When it comes to integrating different individual opinions in group decision making or

advice taking, a typical uncritical strategy is equal weighting of opinions, in spite of better

knowledge or even explicit feedback about clearly unequal competence of different advice givers

(Fiedler et al., 2018; Mahmoodi et al., 2015). Recent research by Powell, Yu, DeWolf, and

Holyoak (2017) showed that the attractiveness of products offered by Amazon may depend on

quantity (number of available reviews) more than on quality (mean rating provided by previous

customers). Confusion of quantity and quality was also observed by Fiedler, Kareev, Avrahami, Beier, Kutzner, and Hütter (2016), who found that increases (decreases) between samples of two symbols in the proportion of one critical symbol were readily detected only when absolute sample size increased (decreased) well.

In causal reasoning, metacognitive myopia is evident in a tendency to exclusively focus on effect strength and to disregard the strength of the causal input that was necessary to induce the observed effect strength. For example, the impact of a drug on athletic performance is judged to be higher if the same dose of the drug causes a performance increase of 10 scale points rather than 1 point. However, whether 254 mg or only 34 mg of the drug were necessary to induce the same observed performance change is given little weight (Hansen, Rim, & Fiedler, 2013).

Why do irrational consequences of metacognitive myopia extend from objectively difficult to such trivially easy task settings? Why do people continue to be influenced by invalid information, which is obviously wrong (like an irrelevant numerical anchor) and which they explicitly classify as invalid? A tentative answer might lie in a kind of metacognitive learned-helplessness effect (Maier & Seligman, 1976). Homo sapiens may have learned that many real-life tasks do not provide us with sufficient information for a normatively sound monitoring and control process. Thus, a Bayesian algorithm required to correct for biases in an information sample is often unknown or does not exist at all. This experience may then be over-generalized to easy situations in which monitoring and control would be simple and straightforward. In any case, metacognitive myopia seems to constitute a major impediment in the way of rational behavior.

Enumerated Summary

1. The term "metacognition" refers to the subset of cognitive operations that are involved in the critical assessment and quality control of one's own cognitive functions. It is useful to distinguish between monitoring and control as the two major metacognitive functions. Rather than being separate from cognition, metacognition is integral to every cognitive performance, from brief perceptual tasks to complex reasoning challenges.

2. Guiding people to effective regulation of effort is the "holy grail", or ultimate goal, of metacognitive research.

3. Crucial for understanding sources for monitoring biases is measurement of *resolution*, the ability to distinguish between correct and incorrect answers, and *calibration*, the extent to which judgments tend to be overconfident or underconfident.

4. Retrospective, post-decision confidence ratings were found to be more accurate than prospective ratings provided beforehand.

5. Metacognitive judgments (e.g., of confidence) utilize distinct heuristic cues, such as fluency, familiarity, accessibility, and pronounceability.

6. Although the neuropsychological underpinnings of metacognition are complicated, convergent evidence indicates that rostral and dorsal parts of the lateral PFC are important for the accuracy of retrospective performance judgments, whereas prospective judgments of performance seem to depend on medial PFC.

7. Metacognitive myopia – the uncritical and naïve tendency to rely on invalid samples of information – constitutes a serious impediment of rational behavior.

Questions for Comprehension and Reflection

1. Is metacognition confined to monitoring and control of conscious and deliberate cognition, or is it also required to regulate automatic processes in low-level cognition?

2. Does metacognition apply to animal learning and decisions?

3. How do we know that we know? Explain the bases of metacognitive judgments.

4. Explain how to measure judgment accuracy and why it is important.

5. How does neuroscience contribute to the understanding of metacognition?

6. What insights and what practical lessons about how to improve learning and test-taking effectively do students gain from an understanding of metacognitive processes?

7. Why is it essential for a teacher, a doctor, and a judge to understand metacognitive processes?

Hot Topics

More on eyewitness memory:

Signal-detection analysis has been extremely helpful in clarifying the metacognitive origin of the serious errors in eyewitness identifications. Even though witness' memory cannot be influenced in retrospect – that is, the discriminability d' of correct and incorrect memories is typically invariant – it has been shown that the rate of false identifications can be markedly reduced by simply inducing a more conservative response strategy, that is, a higher criterion C. A glance at Figure 4 will easily confirm that a rightward shift of C (up to the intersection point of both curves) will reduce the number of incorrect identifications (area right of C under the dashed curve) more than the number of correct identifications (are under the solid curve), thus increasing the overall rate of correct decisions. Such clever metacognition research has led to a commonly noted improvement of legal practices (Wells et al., 2000).

Now after two or three decades of improved lineup procedures, a recent state-of-the art review by Wixted and Wells (2017) has arrived at the optimistic conclusion that "… our understanding of how to properly conduct a lineup has evolved considerably". Under pristine testing conditions (e.g., fair lineups uncontaminated with administrator influence; immediate confidence statement), eyewitness "… (a) confidence and accuracy are strongly related and (b) high-confidence suspect identifications are remarkably accurate." [p. 10].

Computerized learning environments:

Computerized environments are replacing paper-based environments for training, learning, and assessment. However, a puzzling finding is screen inferiority – a disadvantage in learning from computer screens even when the task draws on capabilities considered well-suited for modern technologies like computer or e-books (see Gu, Wu, & Xu, 2015, for a review).

A recently arising metacognitive explanation proposes that computerized environments provide a contextual cue that induces shallower processing than paper environments (e.g., Daniel & Woody, 2013; Morineau, Blanche, Tobin, & Guéguen, 2005). Metacognitive research on reading comprehension has found that JOL reliability is poor (see Dunlosky & Lipko, 2007, for a review). Notably, studies provide growing evidence that associates computerized learning with inferior metacognitive processes, particularly with consistent overconfidence and less effective effort regulation (for a review, see Sidi et al., 2017).

Self-regulated learning needs guidance. Metacognitive scaffolding can support preparatory phases of orientation and planning, monitoring of progress during learning, and retroactive activities such as reflection (Roll, Holmes, Day, & Bonn, 2012). Given that learning and JOL reliability can be improved through self-questioning, appropriate test expectancy, analyzing the task, and delayed summaries (Wiley, Thiede, & Griffin, 2016), it is interesting that screen inferiority could be ameliorated by guiding participants to increase mental effort expenditure. This was achieved by asking participants to proofread, edit, and write keywords summarizing texts' contents (Eden & Eshet-Alkalai, 2013; Lauterman & Ackerman, 2014). Apparently, then, in-depth text processing is the default on paper, whereas on screen an external trigger is required to enhance metacognitive processes leading to enhanced performance (Sidi et al., 2017).

References

Daniel, D. B., & Woody, W. D. (2013). E-textbooks at what cost? Performance and use of electronic v. print texts. *Computers & Education*, *62*, 18–23. https://doi.org/10.1016/j.compedu.2012.10.016

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, *16*(4), 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x

Eden, S., & Eshet-Alkalai, Y. (2013). The effect of format on performance: Editing text in print versus digital formats. *British Journal of Educational Technology*, *44*(5), 846–856.

https://doi.org/10.1111/j.1467-8535.2012.01332.x

Gu, X., Wu, B., & Xu, X. (2015). Design, development, and learning in e-Textbooks: What we learned and where we are going. *Journal of Computers in Education*, *2*(1), 25–41. https://doi.org/10.1007/s40692-014-0023-9

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. *Computers in Human Behavior*, *35*, 455–463. https://doi.org/10.1016/j.chb.2014.02.046

Morineau, T., Blanche, C., Tobin, L., & Guéguen, N. (2005). The emergence of the contextual role of the e-book in cognitive processes through an ecological and functional analysis. *International Journal of Human-Computer Studies*, *62*(3), 329–348. https://doi.org/10.1016/j.ijhcs.2004.10.002

Roll, I., Holmes, N. G., Day, J., & Bonn, D. (2012). Evaluating metacognitive scaffolding in guided invention activities. *Instructional Science*, *40*(4), 691–710. https://doi.org/10.1007/s11251-012-9208-7

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction*, *51*, 61–73. https://doi.org/10.1016/j.learninstruc.2017.01.002

Wells, G. L., Malpass, R. S., Lindsay, R. C., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station. A successful application of eyewitness research. *The American Psychologist*, *55*(6), 581–598. https://doi.org/10.1037/0003-066x.55.6.581

Wiley, J., Thiede, K. W., & Griffin, T. D. (2016). Improving metacomprehension with the situation-model approach. In K. Mokhtari (Ed.), *Improving reading comprehension through metacognitive reading instruction for first and second language readers* (pp. 93–110). Lanham, MD: Rowman & Littlefield.

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65. https://doi.org/10.1177/1529100616686966

Glossary

- **Monitoring**: Critical assessment of one's ongoing cognitive activities in light of the goals of the task. Results of the monitoring function trigger the complementary **control** function, which entails corrections and strategic changes to ensure that goals are achieved more effectively or efficiently.

- **Control and effort regulation**: Adjustment of one's cognitive or behavioral responses, triggered by a monitoring process, in order to achieve higher levels of performance accuracy. Control can take many different forms depending on the task at hand and the available resources.

- **Meta-Memory**: Monitoring and control of memorizing and retrieval processes.

- **Meta-Reasoning**: Monitoring and control processes regarding problem solving, reasoning, and decision-making processes.

- **Confidence**: Subjective feeling that one' decision or response is correct, typically elicited as a probability judgment.

- **Judgment of learning (JOL)**: Prediction of the likelihood of recall for recently studied items. JOL has an essential role in study time allocation and effort regulation.

- **Resolution**: Extent in which confidence judgments regarding several tasks of the same kind (e.g., answers in an exam) indeed distinguish between correct and incorrect responses

- **Calibration and overconfidence**: Calibration is the gap between actual chances of success and confidence judgments of success. An overconfidence illusion reflects the inflated confidence in one's actual performance, ability, or chance of success.

- **Metacognitive heuristic cues**: Since individuals do not have direct access to their knowledge and the quality of their own cognitive processing, metacognitive judgments are inferred from heuristic cues experienced during task performance and immediately afterword.

- **Metacognitive myopia**: Phenomenon by which people are pretty accurate in utilizing

  information given in a sample, whereas they are naive and almost blind regarding the validity

  of the sample.

- **Neuroimaging**: use of various tools to produce images of the nervous system. Structural

  imaging (e.g., MRI) provides images of the brain's anatomical structure, whereas functional

  imaging (e.g., fMRI) provides images of the brain as individuals complete mental tasks. The

  brain areas involved in responding to current tasks "light up", mapping the parts of the brain

  involved in a given task.

- **Prospective and retrospective judgments**: In metacognitive research, subjects are explicitly

  asked to make judgments about their performance on a criterion task. These judgments can be

  prospective, and ask subjects to judge their future performance, or they may be retrospective,

  and ask subjects to judge their prior performance.

References

Ackerman, R. (2014). The diminishing criterion model for metacognitive regulation of time investment. *Journal of Experimental Psychology: General*, *143*(3), 1349–1368. https://doi.org/10.1037/a0035098

Ackerman, R., & Beller, Y. (2017). Shared and distinct cue utilization for metacognitive judgments during reasoning and memorization. *Thinking & Reasoning*, *23*(4), 376–408. https://doi.org/10.1080/13546783.2017.1328373

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, *17*(1), 18–32. https://doi.org/10.1037/a0022086

Ackerman, R., & Thompson, V. A. (2015). Meta-Reasoning: What can we learn from meta-memory? In A. Feeney & V. Thompson (Eds.), *Reasoning as Memory* (pp. 164–182). Hove, UK: Psychology Press.

Ackerman, R., & Thompson, V. A. (2017). Meta-Reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences*, *21*(8), 607–617. https://doi.org/10.1016/j.tics.2017.05.004

Ackerman, R., & Zalmanov, H. (2012). The persistence of the fluency–confidence association in problem solving. *Psychonomic Bulletin & Review*, *19*(6), 1187–1192. https://doi.org/10.3758/s13423-012-0305-z

Ahrendt, H. (1963). *Eichmann in Jerusalem: A report on the banality of evil*. London, UK: Faber & Faber.

Ais, J., Zylberberg, A., Barttfeld, P., & Sigman, M. (2016). Individual consistency in the accuracy and distribution of confidence judgments. *Cognition*, *146*, 377–386. https://doi.org/10.1016/j.cognition.2015.10.006

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science*, *329*(5995), 1081–1085. https://doi.org/10.1126/science.1185718

Boldt, A., De Gardelle, V., & Yeung, N. (2017). The impact of evidence reliability on sensitivity and bias in decision confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(8), 1520–1531. https://doi.org/10.1037/xhp0000404

Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, *28*(11), 1531–1546. https://doi.org/10.1177/0956797617714579

Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009). Neural correlates of metamemory: A comparison of feeling-of-knowing and retrospective confidence judgments. *Journal of Cognitive Neuroscience*, *21*(9), 1751–1765. https://doi.org/10.1162/jocn.2009.21123

Danion, J. M., Gokalsing, E., Robert, P., Massin-Krauss, M., & Bacon, E. (2001). Defective relationship between subjective experience and behavior in schizophrenia. *American Journal of Psychiatry*, *158*(12), 2064–2066. https://doi.org/10.1176/appi.ajp.158.12.2064

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science, 5*(5), 303–305. https://doi.org/10.1111/j.1467-9280.1994.tb00630.x

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, (25), 23–38. https://doi.org/10.1016/j.edurev.2018.09.003

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, *29*(5), 761–778. https://doi.org/10.1177/0956797617744771

Do Lam, A. T. A., Axmacher, N., Fell, J., Staresina, B. P., & Gauggel, S. (2012). Monitoring the mind: The neurocognitive correlates of metamemory. *PLoS ONE*, *7*(1), e30009. https://doi.org/10.1371/journal.pone.0030009

Dunlosky, J., & Ariel, R. (2011). The influence of agenda-based and habitual processes on item selection during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 899–912. https://doi.org/10.1037/a0023064

Dunlosky, J., & Tauber, S. K. (2014). Understanding people's metacognitive judgments: An isomechanism framework and its implications for applied and theoretical research. In T. Perfect (Ed.), *Handbook of applied memory* (pp. 444–464). Thousand Oaks, CA: Sage.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Fernandez-Cruz, A. L., Arango-Muñoz, S., & Volz, K. G. (2016). Oops, scratch that! Monitoring one's own errors during mental calculation. *Cognition*, *146*, 110–120. https://doi.org/10.1016/j.cognition.2015.09.005

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, *107*(4), 659–676. https://doi.org/10.1037/0033-295x.107.4.659

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 186–203. https://doi.org/10.1037/0278-7393.34.1.186

Fiedler, K. (2012). Meta-cognitive myopia and the dilemmas of inductive-statistical inference. In B. H. Ross & B. H. Ross (Eds.), *The psychology of learning and motivation (Vol 57)* (pp. 1–55). San Diego, CA: Elsevier Academic Press. https://doi.org/10.1016/b978-0-12-394293-7.00001-7

Fiedler, K. (2013). Fluency and behavior regulation: Adaptive and maladaptive consequences of a good feeling. In C. Unkelbach & R. Greifender (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behaviour* (pp. 234–254). New York, NY: Psychology Press.

Fiedler, K., Armbruster, T., Nickel, S., Walther, E., & Asbeck, J. (1996). Constructive biases in

social judgment: Experiments on the self-verification of question contents. *Journal of Personality and Social Psychology*, *71*(5), 861–873. https://doi.org/10.1037/0022-3514.71.5.861

Fiedler, K., Hütter, M., Schott, M., & Kutzner, F. (2018). Metacognitive myopia and the overutilization of misleading advice. *Journal of Behavioral Decision Making*, Advance online publication. https://doi.org/10.1002/bdm.2109

Fiedler, K., Kareev, Y., Avrahami, J., Beier, S., Kutzner, F., & Hütter, M. (2016). Anomalies in the detection of change: When changes in sample size are mistaken for changes in proportions. *Memory & Cognition*, *44*(1), 143–161. https://doi.org/10.3758/s13421-015-0537-z

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1594), 1338–1349. https://doi.org/10.1098/rstb.2011.0417

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *Journal of Neuroscience*, *32*(18), 6117–6125. https://doi.org/10.1523/jneurosci.6489-11.2012

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, 443. https://doi.org/10.3389/fnhum.2014.00443

Fleming, S. M., Massoni, S., Gajdos, T., & Vergnaud, J.-C. (2016). Metacognition about the past and future: Quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness*, *2016*(1), 1–12. https://doi.org/10.1093/nc/niw018

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain : A Journal of Neurology*, *137*(10), 2811–2822. https://doi.org/10.1093/brain/awu221

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), 1541–1543. https://doi.org/10.1126/science.1191883

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*(4), 506–528. https://doi.org/10.1037/0033-295x.98.4.506

Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, *32*(1), 1–36. https://doi.org/10.1007/s10713-007-0003-3

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, UK: John Wiley.

Hansen, J., Rim, S., & Fiedler, K. (2013). Psychological distance and judgments of causal impact. *Journal of Experimental Social Psychology*, *49*(6), 1184–1189. https://doi.org/10.1016/j.jesp.2013.05.012

Hembacher, E., & Ghetti, S. (2017). Subjective experience guides betting decisions beyond accuracy: Evidence from a metamemory illusion. *Memory, 25*(5), 575–585. https://doi.org/10.1080/09658211.2016.1197946

Higham, P. A., & Higham, D. P. (2018). New improved gamma: Enhancing the accuracy of Goodman–Kruskal's gamma using ROC curves. *Behavior Research Methods*, Advance online publication. https://doi.org/10.3758/s13428-018-1125-5

Jackson, S. A., & Kleitman, S. (2014). Individual differences in decision-making and confidence: Capturing decision tendencies in a fictitious medical test. *Metacognition and Learning, 9*(1), 25–49. https://doi.org/10.1007/s11409-013-9110-y

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, 35*(2), 157–175. https://doi.org/10.1006/jmla.1996.0009

Koren, D., Seidman, L. J., Poyurovsky, M., Goldsmith, M., Viksman, P., Zichel, S., & Klein, E. (2004). The neuropsychological basis of insight in first-episode schizophrenia: A pilot metacognitive study. *Schizophrenia Research, 70*(2–3), 195–202. https://doi.org/10.1016/j.schres.2004.02.004

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*(4), 609–639. https://doi.org/10.1037/0033-295x.100.4.609

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A. (2012). When are two heads better than one and why? *Science, 336*(6079), 360–362. https://doi.org/10.1126/science.1216549

Koriat, A. (2018). When reality is out of focus: Can people tell whether their beliefs and judgments are correct or wrong? *Journal of Experimental Psychology: General, 147*(5), 613–631. https://doi.org/10.1037/xge0000397

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133*(4), 643–656. https://doi.org/10.1037/0096-3445.133.4.643

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 34–53. https://doi.org/10.1037/0278-7393.27.1.34

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135*(1), 36–68. https://doi.org/10.1037/0096-3445.135.1.36

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review, 14*(2), 219–224. https://doi.org/10.3758/bf03194055

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

Lynch., P. J., & Jaffe, C. C. (2006). *Brain human lateral view*. Retrieved from https://commons.wikimedia.org/wiki/File:Brain_human_lateral_view.svg

Lynch., P. J., & Jaffe, C. C. (2006). *Brain human sagittal section*. Retrieved from https://commons.wikimedia.org/wiki/File:Brain_human_sagittal_section.svg

Mahmoodi, A., Bang, D., Olsen, K., Zhao, Y. A., Shi, Z., Broberg, K., … Bahrami, B. (2015). Equality bias impairs collective decision-making across cultures. *Proceedings of the National Academy of Sciences*, *112*(12), 3835–3840. https://doi.org/10.1073/pnas.1421692112

Maier, S. F., & Seligman, M. E. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, *105*(1), 3–46. https://doi.org/10.1037/0096-3445.105.1.3

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*(1), 174–179. https://doi.org/10.3758/pbr.15.1.174

Metcalfe, J., & Schwartz, B. L. (2016). The ghost in the machine: Self-reflective consciousness and the neuroscience of metacognition. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 407–424). Oxford, UK: Oxford University Press.

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *Journal of Neuroscience*, *38*(14), 3534–3546. https://doi.org/10.1523/jneurosci.2360-17.2018

Mueller, M. L., & Dunlosky, J. (2017). How beliefs can impact judgments of learning: Evaluating analytic processing theory with beliefs about fluency. *Journal of Memory and Language*, *93*, 245–258. https://doi.org/10.1016/j.jml.2016.10.008

Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2013). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review*, *20*(2), 378–384. https://doi.org/10.3758/s13423-012-0343-6

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*(1), 109–133. https://doi.org/10.1037/0033-2909.95.1.109

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation*, *26*, 125–173. https://doi.org/10.1016/s0079-7421(08)60053-5

Nguyen, T. B., Abed, E., & Pezdek, K. (2018). Postdictive confidence (but not predictive confidence) predicts eyewitness memory accuracy. *Cognitive Research: Principles and Implications*, *3*(32), 1–13. https://doi.org/10.1186/s41235-018-0125-4

Pannu, J. K., Kaszniak, A. W., & Rapcsak, S. Z. (2005). Metamemory for faces following frontal lobe damage. *Journal of the International Neuropsychological Society*, *11*(6), 668–676. https://doi.org/10.1017/s1355617705050873

Powell, D., Yu, J., DeWolf, M., & Holyoak, K. J. (2017). The love of large numbers: A popularity bias in consumer choice. *Psychological Science*, *28*(10), 1432–1442. https://doi.org/10.1177/0956797617711291

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 435–451. https://doi.org/10.1037/0278-7393.18.3.435

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615–625. https://doi.org/10.1037/a0013684

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, *137*(1), 131–148. https://doi.org/10.1037/a0021705

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, *32*(5), 880–892. https://doi.org/10.1037/0022-3514.32.5.880

Rouault, M., Mcwilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: Insights from individual differences and neuroimaging. *Personality Neuroscience*, *1*, e17. https://doi.org/10.1017/pen.2018.16

Schnyer, D. M., Verfaellie, M., Alexander, M. P., LaFleche, G., Nicholls, L., & Kaszniak, A. W. (2004). A role for right medial prefrontal cortex in accurate feeling-of-knowing judgments: Evidence from patients with lesions to frontal cortex. *Neuropsychologia*, *42*(7), 957–966. https://doi.org/10.1016/j.neuropsychologia.2003.11.020

Shanks, L. L., & Serra, M. J. (2014). Domain familiarity as a cue for judgments of learning. *Psychonomic Bulletin & Review*, *21*(2), 445–453. https://doi.org/10.3758/s13423-013-0513-1

Shimamura, A. P., & Squire, L. R. (1986). Memory and metamemory: A study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *12*(3), 452–460. https://doi.org/10.1037/0278-7393.12.3.452

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction*, *51*, 61–73. https://doi.org/10.1016/j.learninstruc.2017.01.002

Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I was so sure! Metacognitive judgments are less accurate given prospectively than retrospectively. *Frontiers in Psychology*, *7*(218), 1–8. https://doi.org/10.3389/fpsyg.2016.00218

Son, L. K., & Sethi, R. (2010). Adaptive learning and the allocation of time. *Adaptive Behavior*, *18*(2), 132–140. https://doi.org/10.1177/1059712309344776

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York, NY: Doubleday.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*(1), 66–73. https://doi.org/10.1037/0022-0663.95.1.66

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 1024–1037. https://doi.org/10.1037/0278-7393.25.4.1024

Thompson, V. A., Prowse Turner, J. A., Pennycook, G., Ball, L., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, *128*, 237–251. https://doi.org/10.1016/j.cognition.2012.09.012

Topolinski, S., Bakhtiari, G., & Erle, T. M. (2016). Can I cut the Gordian tnok? The impact of pronounceability, actual solvability, and length on intuitive problem assessments of anagrams. *Cognition*, *146*, 439–452. https://doi.org/10.1016/j.cognition.2015.10.019

Topolinski, S., & Reber, R. (2010). Immediate truth-Temporal contiguity between a cognitive problem and its solution determines experienced veracity of the solution. *Cognition*, *114*(1), 117–122. https://doi.org/10.1016/j.cognition.2009.09.009

Undorf, M., & Ackerman, R. (2017). The puzzle of study time allocation for the most challenging items. *Psychonomic Bulletin & Review*, *24*(6), 2003–2011. https://doi.org/10.3758/s13423-017-1261-4

Undorf, M., & Erdfelder, E. (2015). The relatedness effect on judgments of learning: A closer look at the contribution of processing fluency. *Memory & Cognition*, *43*(4), 647–658. https://doi.org/10.3758/s13421-014-0479-x

Undorf, M., & Zimdahl, M. F. (2018). Metamemory and memory for a wide range of font sizes: What is the contribution of perceptual fluency? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *45*(1), 97–109. https://doi.org/10.1037/xlm0000571

Undorf, M., Zimdahl, M. F., & Bernstein, D. M. (2017). Perceptual fluency contributes to effects of stimulus size on judgments of learning. *Journal of Memory and Language*, *92*, 293–304. https://doi.org/10.1016/j.jml.2016.07.003

Unkelbach, C., Fiedler, K., & Freytag, P. (2007). Information repetition in evaluative judgments: Easy to monitor, hard to control. *Organizational Behavior and Human Decision Processes*, *103*(1), 37–52. https://doi.org/10.1016/j.obhdp.2006.12.002

Vaccaro, A. G., & Fleming, S. M. (2018). Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain and Neuroscience Advances*, *2*, 1–14. https://doi.org/10.1177/2398212818810591

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring

effects: basic anchoring and its antecedents. *Journal of Experimental Psychology. General*, *125*(4), 387–402. https://doi.org/10.1037/0096-3445.125.4.387

Zacks, R. T. (1969). Invariance of total learning time under different conditions of practice. *Journal of Experimental Psychology*, *82*(3), 441–447. https://doi.org/10.1037/h0028369