Oxford Handbooks Online

The Self-Consistency Theory of Subjective Confidence

Asher Koriat and Shiri Adiv

The Oxford Handbook of Metamemory *(Forthcoming)*
Edited by John Dunlosky and Sarah (Uma) K. Tauber

Online Publication Date: Jun 2015
Subject: Psychology, Cognitive Psychology, Clinical Psychology
DOI: 10.1093/oxfordhb/9780199336746.013.18

Abstract and Keywords

Innumerable studies have yielded a positive correlation between subjective confidence and accuracy, suggesting that people are skillful in discriminating between correct and wrong answers. The chapter reviews evidence from different domains indicating that people's subjective confidence in an answer is diagnostic of the consensuality of the answer rather than of its accuracy. A self-consistency model (SCM) was proposed to explain why the confidence-accuracy correlation is positive when the correct answer is the consensually chosen answer but is negative when the wrong answer is the consensual answer. Several results that were obtained across a variety of tasks provided support for the generality of the theoretical framework underlying SCM.

Keywords: subjective confidence, confidence-accuracy relationship, self-consistency model, consensuality principle, Wisdom of crowds, overconfidence

When people are asked to answer a question or to solve a problem, they can indicate their confidence that the answer or solution is correct. Confidence judgments have been used and investigated in a wide range of domains. These domains include perception and psychophysics, memory and metacognition, decision-making and choice, eyewitness testimony, scholastic achievement and intelligence, social cognition, neuroscience, and animal cognition. Of course, philosophers have also been concerned with the issue of how we can be sure about the truth of assertions (e.g., BonJour, 1985; Engel, 1998). Statisticians also examined these questions from a normative perspective, focusing on the degree of confidence in conclusions that are based on empirical observations (Fisher, 1925; Lykken, 1968).

In experimental settings, the collection of confidence judgments was used for different goals. In perception and psychophysics, confidence judgments have been used to explore different quantitative theories of the processes underlying psychophysical judgments (Vickers, Smith, Burt, & Brown, 1985; Wixted, & Mickes, 2010). Forensic psychologists have focused primarily on questions regarding the validity of confidence as a diagnostic cue of the accuracy of a testimony (Bothwell, Deffenbacher, & Brigham, 1987; Read, Lindsay, & Nicholls, 1998; Sporer, Penrod, Read, & Cutler, 1995). Among social psychologists and memory researchers, confidence judgments have attracted attention specifically because these judgments have been found to moderate the likelihood of translating one's beliefs into behavior (Ross, 1997; Tormala & Rucker, 2007; Yzerbyt, Lories, & Dardenne, 1998). Vickers (2001), however, complained that "the variable of confidence seems to have played a Cinderella role in cognitive psychology—relied on for its usefulness, but overlooked as an interesting variable in its own right" (p. 148). Fortunately, there has been increased interest in the study of subjective confidence in its own right, including the processes underlying confidence judgments, and the determinants of their accuracy and inaccuracy.

Core Questions

Three issues constitute the core issues in metacognition research on subjective confidence. The first concerns the correspondence between confidence and performance: How faithful are confidence judgments in mirroring object-level performance? Second, what are the processes underlying the subjective feeling of certainty and doubt? Finally, given the postulated bases of confidence, how do these bases explain the accuracy and inaccuracy of confidence judgments under different conditions?

The accuracy of confidence judgments.

The first question, which has received a great deal of research, concerns the accuracy of confidence judgments (e.g., Dunning, Heath, & Suls, 2004; Liberman & Tversky, 1993; Lichtenstein, Fischhoff, & Phillips, 1982). Researchers in the area of judgment and decision-making (see Lichtenstein et al., 1982; Murphy, 1973) have provided a methodology for deriving different scores that convey information about two aspects of metacognitive accuracy, calibration and resolution (see Dunlosky, Mueller, & Thiede, this volume; Higham, Zawadzka, & Hanczakowski, this volume(. Calibration (“bias”) or “absolute accuracy” (see Nelson & Dunlosky, 1991) refers roughly to the correspondence between mean metacognitive judgments and mean actual performance, and reflects the extent to which confidence judgments are realistic or disclose overconfidence bias (inflated confidence relative to performance) or underconfidence bias. Calibration can be evaluated only when judgments and performance are measured on equivalent scales. Such is not the case for the second aspect of metacognitive accuracy, resolution (or relative accuracy). Resolution refers to the extent to which metacognitive judgments are correlated with memory performance across items. This aspect, which is commonly indexed by a within-subject gamma correlation between judgments and performance (Nelson, 1984) reflects the ability to discriminate between correct and incorrect answers. Researchers in the area of judgment and decision-making have provided a methodology for deriving different scores from calibration curves (see Lichtenstein et al., 1982; Murphy, 1973).

There has been a division of labor in studies of confidence accuracy such that the work within the judgment and decision tradition has focused on the calibration of subjective probabilities (e.g., Griffin & Brenner, 2004). In contrast, the work in metacognition by cognitive psychologists has focused primarily on resolution—the discrimination between correct and wrong answers or judgments (see Koriat, 2007; Metcalfe & Dunlosky, 2008). The observation that people can tell when they are right and when they are wrong has been among the steering forces for the upsurge of interest in metacognition (Hart, 1965; Koriat, 1993; Nelson & Dunlosky, 1991; Tulving & Madigan, 1970). Surprisingly, this observation has received relatively little attention among students of judgment and decision-making despite the fact that virtually all calibration curves reported in the experimental literature are monotonically increasing, suggesting good resolution (see Keren, 1991). In fact, in studies of recognition memory, it has been noted that low-confidence decisions are associated with close-to-chance accuracy, whereas high-confidence decisions tend to be associated with close-to-perfect accuracy (Mickes, Hwe, Wais, & Wixted, 2011). Similarly, studies that collected confidence judgments in a variety of tasks, have generally yielded moderate to high within-person confidence/accuracy (C/A) across items (e.g., Brewer, Keast, & Rishworth, 2002; Lindsay, Wells, & Rumpel, 1981). However, the extensive research on “assessed probabilities” has focused on patterns of miscalibration (e.g., Griffin & Brenner, 2004), taking for granted the accuracy of monitoring resolution. Furthermore, within the judgment and decision tradition there seems to have been an implicit assumption that assessed probabilities ought to be perfectly calibrated, and hence the challenge is to explain deviations from perfect calibration. In metacognition research, in contrast, one of the research goals has been to uncover the bases of confidence judgments and to explain why these judgments are largely accurate.

The bases of confidence judgments.

The second question about confidence judgments concerns their bases. Three theoretical approaches to this question have been distinguished: the *direct-access* approach, the *information-based* approach, and the *experience-based* approach (Koriat, 1997; Koriat & Levy-Sadot, 1999). Research on the possible bases of confidence judgments seems to have been hampered by the implicit endorsement of the direct-access view according to which these judgments mirror directly memory strength. For example, in strength theories of memory, the assumption is that confidence judgments are scaled from the strength or quality of the internal memory representation (see Van Zandt, 2000, for a review). The direct-access view has resulted in taking the accuracy of confidence judgments for granted.

The direct-access approach is perhaps best represented in the philosophy of knowledge by the claims of rationalist philosophers that a priori truths (e.g., mathematical propositions) are based on intuition and deduction, and that their certainty is self-evident (see Koriat, 2012b). This approach seems to find some expression in the experimental literature on confidence in the idea that some answers and their associated strong confidence are based on "direct retrieval" (Juslin, Winman, & Olsson, 2003; Metcalfe, 2000; Unkelbach & Stahl, 2009). Whether the response to such questions (e.g., "what is your name?" see Koriat, 2012b) should be assumed to involve a non-inferential basis is still an open question.

In contrast to the direct-access view, most researchers in metacognition lean towards the assumption that metacognitive judgments are inferential in nature, relying on a variety of beliefs and heuristics that may be applied under different conditions (see Benjamin & Bjork, 1996; Koriat, 1997; Koriat, Ma'ayan, & Nussinson, 2006). A distinction is drawn, however, between information-based and experience-based judgments (Kelley & Jacoby, 1996; Koriat, Nussinson, Bless, & Shaked, 2008). Information-based judgments rely on an analytic inference in which various considerations are consulted to reach an educated judgment. This view is in the spirit of the reason-based approach of Shafir, Simonson and Tversky (1993). They argued that when faced with the need to choose, people often seek and construct reasons in order to resolve the conflict and justify their choice. For example, confidence in two-alternative forced-choice (2AFC) general-knowledge questions was claimed to rest on the reasons recruited in favor of the two answers (e.g., Griffin & Tversky, 1992; Koriat, Lichtenstein, & Fischhoff, 1980; McKenzie, 1997).

Experience-based judgments, in contrast, are based on mnemonic cues that derive on-line from task performance rather than being based on the content of domain-specific declarative information retrieved from long-term memory. For example, confidence judgments are said to rest on the ease with which information comes to mind or on the speed with which an answer is selected among distractors (e.g., Kelley & Lindsay, 1993; Koriat et al., 2006; Robinson, Johnson, & Herndon, 1997). The heuristics that shape subjective confidence are assumed to operate largely below full consciousness (Koriat, 2000).

The reasons for the accuracy and inaccuracy of confidence judgments.

The third question concerns the processes underlying the accuracy and inaccuracy of confidence judgments. As noted earlier, the direct-access view to confidence judgments take the accuracy of these judgments for granted to the extent that confidence judgments are assumed to convey information about object-level processes. Inferential approaches to confidence, in contrast, are faced with the challenge of explaining the accuracy of confidence judgments.

There has been some debate in the literature regarding the validity of some of the findings documenting systematic discrepancies between confidence and performance. On the one hand, proponents of the ecological probability approach (Dhami, Hertwig, & Hoffrage, 2004; Gigerenzer, Hoffrage, & Kleinbölting, 1991) argued that some of these discrepancies are artifactual, deriving from the failure of researchers to follow the dictum of a representative research design (Brunswik, 1956). Thus, they argued that the overconfidence bias (Hoffrage, 2004) and the hard-easy effect (Griffin & Tversky, 1992) that had been observed in studies of confidence stem from researchers' failure to sample items so that they are representative of the natural environment. Indeed, several studies that used a set of items that were randomly selected from a circumscribed domain of knowledge found little evidence for overconfidence bias or for the hard-easy effect (Gigerenzer et al., 1991; Juslin, 1993, 1994). On the other hand, among researchers in metacognition, the cue-utilization view has led to a deliberate focus on the inaccuracies of metacognitive judgments, in general, and confidence judgments, in particular. A large number of studies documented systematic discrepancies between subjective and objective indexes of knowledge.

Koriat, Pansky, and Goldsmith (2011) argued that the difference between the two lines of research, one emphasizing a representative design, and another focusing on metacognitive illusions, reflect a difference in research agendas. The first agenda is to obtain a faithful description of the state of affairs in the real world. This agenda requires that the experimental conditions should be representative of conditions and variations in the real world. The second agenda is to achieve a theoretical understanding of the phenomena and their underlying mechanisms. This agenda, in contrast, sometimes calls precisely for the use of conditions that are ecologically unrepresentative, even contrived, in order to untangle variables that go hand in hand in real life (see Koriat, 2012a). Indeed, in metacognition research, researchers have sometimes deliberately focused on factors that lead metacognitive judgments astray (e.g., Benjamin, Bjork, & Schwartz, 1998; Brewer & Sampaio, 2006; Busey, Tunnicliff, Loftus, & Loftus, 2000; Chandler, 1994; Koriat, 1995; Rhodes & Castel, 2008).

The Motivation for the Present Proposal

The motivation for the self-consistency model of subjective confidence derived initially from the results of an old study (Koriat, 1975) that examined the C/A relationship in a phonetic symbolism task. In previous studies (e.g., Slobin, 1968), participants were asked to match antonymic pairs from noncognate languages (e.g., *tuun-luk*) with their English equivalents (*deep-shallow)*. The results indicated that people's matches are significantly better than chance. Koriat (1975) examined whether participants can also monitor the correctness of their matches, and asked participants to indicate their confidence in each match. Participants' object-level accuracy was significantly better than chance: Participants' matches were correct in 58% of the cases. In addition, their meta-level accuracy was also significant: The percentage of correct matches increased steeply with confidence judgments, suggesting that participants were successful in monitoring the correctness of their matches. The latter result presented a puzzle. Neither the information-based approach nor the experience-based approach offers a hint regarding the cues that participants might use to monitor the correctness of their matches. The finding is reminiscent of the direct-access view that rationalists posit with regard to a priori propositions that are accessed through intuition.

In an attempt to explain the high C/A correlation, Koriat (1976, see Study 1 in Table 1) suggested that perhaps the observation that participants' matches are largely accurate ("knowledge") might create a confounding for the assessment of the C/A correlation ("metaknowledge"). That is, the correct match is the one that is consensually endorsed, so confidence judgments might actually be correlated with the consensuality of the match rather than with its correctness. Indeed, the results of a subsequent study (Koriat, 1976) confirmed that possibility. In that study, a deliberate effort was made to include a large proportion of items for which participants would be likely to agree on the wrong match. The items were classified post-hoc into three classes according to whether the majority of participants agreed on the correct match (consensually-correct; CC), agreed on the wrong match (consensually-wrong; CW), or did not agree on either match (nonconsensual; NC). The results clearly indicated that confidence judgments correlated with the consensuality of the match rather than with its correctness: For the CC class, correct matches were endorsed with stronger confidence than were wrong matches, whereas for the CW class, wrong matches were actually associated with stronger confidence than were correct matches. For the NC class, confidence was unrelated to the correctness of the match.

This interactive pattern was referred to as the *consensuality principle* (Koriat, 2008). This pattern was found to be true for several domains, as will be reviewed. The results suggest that the positive C/A correlation that has been observed in a great number of studies is actually because in practically all of these studies participant were more often correct than wrong (i.e., the great majority of items are CC items). Consider, for example, studies of confidence judgments in 2AFC general-information questions. Participants' proportion of correct answers is typically well above .50, and rarely does any of the questions yield more wrong answers than correct answers. The latter questions were sometimes referred to as "deceptive." "misleading," or "unrepresentative" (Fischhoff, Slovic, & Lichtenstein, 1977; Gigerenzer, et al., 1991). Similarly, in psychophysical experiments, judgments tend to be largely accurate with the exception of occasional errors that are not correlated across participants (see Juslin & Olsson, 1997). As a result, the C/A correlation for such questions is typically assessed only across half of the range of proportion correct (.51–1.00), and the range between 0 and .50 is hardly represented.

Before we describe the model, we should say a few words about the methodology of the studies on which it was based. In each of these studies, participants answered a series of 2AFC questions. For each question, they chose one answer and indicated their confidence. As noted earlier, SCM was initially motivated by attempts to clarify the accuracy of confidence judgments. However, the results led to the question of the basis of these judgments. Because this question applies also to domains in which the response does not have a truth-value, SCM was extended to the investigation of the process underlying confidence judgments in such domains as social attitudes and social beliefs, personal preferences, and the category membership decisions.

The Self-Consistency Model of Subjective Confidence

SCM adopts the metaphor of an intuitive statistician underlying human decision and choice (Peterson & Beach, 1967; see McKenzie, 2005). It assumes that the process underlying choice and confidence is analogous to that in which information is sampled from the outside world with the intention (a) to test a hypothesis about a population and (b) to assess the likelihood that the conclusion reached is correct.

It was proposed that when presented with a 2AFC item, it is by replicating the choice process several times that one can appreciate the degree of doubt or certainty involved. Subjective confidence is based on the consistency with which different replications agree in favoring one of the two choices. It represents essentially an assessment of *reproducibility*—the likelihood that a new replication of the decision process will yield the same choice. Thus, *reliability* is used by participants as a cue for *validity*. This is very much like statistical inference when conclusions about a population are based on a sample of observations drawn from that population (Koriat, 2012a).

Thus, SCM incorporates a sampling assumption that is common in many decision models (e.g., Juslin & Olsson, 1997; Stewart, 2009; Stewart, Chater, & Brown, 2006; Vickers & Pietsch, 2001; Vul, Goodman, Griffiths, & Tenenbaum, 2009). When presented with a 2AFC item, participants are assumed to sample a number of representations from a population of potential representations associated with the item. The term “representation” was used as an abstract term that may apply to different 2AFC tasks. It may include a specific consideration (Koriat, Lichtenstein, & Fischhoff, 1980), a particular interpretation or framing of a choice problem (Tversky & Kahneman, 1981), a “cue” that is used to infer the answer (Gigerenzer et al., 1991), or any hunch or association that may tip the balance in favor of one choice rather than the other. Because of the limitations of the cognitive system, the number of representations sampled on each occasion must be quite limited, because of the need to integrate information across representations.

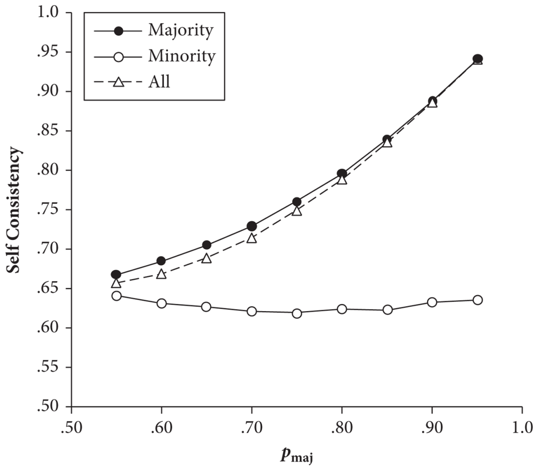
Participants are assumed to draw the implications of each representation, and reach an ultimate decision based on the balance of evidence in favor of the two options (Vickers, 2001; see Baranski & Petrusic, 1998). Once a choice has been made, confidence is based primarily on self-consistency—the general agreement among the sampled representations in favoring the decision reached. In SCM, self-consistency is conceptualized as a contentless cue that reflects the mere number of pro and con considerations associated with the choice irrespective of their meaning and importance (see Alba & Marmorstein, 1987). Clearly, the type of representations retrieved in making a choice should differ depending on the domain of the question. However, SCM assumes that the gross architecture of the process is similar across a variety of 2AFC tasks.

An important assumption of SCM is that in responding to 2AFC items, whether they involve general-information questions or beliefs and attitudes, participants with the same experience draw representations largely from the same, commonly shared population of representations associated with each item. Thus, although the specific samples drawn on each occasion may differ for different individuals and for each individual on different occasions, people draw their clues from a pool of clues that is largely commonly shared. In the case of general-information and perceptual judgments proponents of the ecological approach to cognition (Dhami et al., 2004; Gigerenzer, 2008; Juslin, 1994) have stressed the general accuracy of the shared knowledge, which is assumed to result from the adaptation to the natural environment. In addition, the *wisdom-of-crowds* phenomenon suggests that information that is aggregated across participants is generally closer to the truth than the information provided by each individual participant (Galton, 1907; Mozer, Pashler, & Homaei, 2008; Wallsten, Budescu, Erev, & Diederich, 1997). Thus, we assume that the ingredients that participants use to construct their decisions are drawn from a collective “wisdom.” This is the reason for the observation that confidence judgments are diagnostic of the consensuality of the choice.

Implementation of SCM for the Basis of Confidence Judgments

In what follows, we present a specific instantiation of the model that is clearly oversimple but is sufficient for bringing to the fore the main predictions of SCM. In this instantiation we assume the following: (1) For each 2AFC item, a maximum number of representations (n_{max}) is sampled randomly. (2) Each representation yields a binary subdecision, favoring one of the two options. (3) When a sequence of a preset number (n_{run}) of representations yields the same subdecision, the sampling is stopped, and that subdecision dictates the choice (see Audley, 1960). (4) Each subdecision makes an equal contribution to the ultimate, overt decision and to a self-consistency index, which is assumed to underlie subjective confidence.

To examine the implications of the model, a simulation experiment was run (see Koriat, 2012a; Koriat & Adiv, 2011) in which n_{max} was set at 7. Also, n_{run} was set at 3, so that the actual size of the sample (n_{act}) underlying choice and confidence could vary between 3 and 7. Assume that each item is characterized by a probability distribution, with p_{maj} denoting the probability that a representation favoring the majority choice will be sampled. This probability can be seen as a property of a binary choice item.



Click to view larger

Figure 1. Self-consistency scores as a function of the probability of drawing a majority representation (*Pmaj*) based on the results of the simulation experiment. Reproduced with permission from Koriat and Adiv (2011). Copyright 2011 by Guilford Press.

A simulation experiment was run which assumed nine binomial populations that differ in p_{maj}, with p_{maj} varying from .55 to .95, at .05 steps. For each population, 90,000 iterations were run in each of which a sample of (3-7) representations was drawn. The ultimate choice was classified as "majority" when it corresponded to the majority value in the population (the one that is consistent with p_{maj}), and as "minority" when it corresponded to the minority value in the population. A self-consistency index was calculated for each iteration, which is inversely related to the sample standard deviation. It was defined as $1 - \sqrt{\hat{p}\hat{q}}$ (range .5–1.0), when *p* and *q* designate the proportion of representations favoring the two choices, respectively.

Based on the results of the simulation, Figure 1 presents the self-consistency index, which is assumed to underlie subjective confidence, for majority and minority choices and for all choices combined as a function of p_{maj}. Self-consistency increases monotonically with p_{maj}, but more important, self-consistency is higher for majority than for minority choices. This is because as long as p_{maj} >.50, majority choices will be supported by a larger proportion of the sampled representations than minority choices. For example, for p_{maj} = .70, and sample size = 7, the likelihood that six or seven representations will favor the majority choice is .329, whereas only in .004 of the samples will six or seven representations favor the minority choice. Thus, the expectation is that confidence should be higher for majority choices than for minority choices.

Of course, p_{maj} for a particular item is not known. However, it can be estimated from pc_{maj}—the probability with which the majority alternative is chosen. The theoretical function relating pc_{maj} to p_{maj} can be obtained from the simulation just described. pc_{maj} is an accelerated function of p_{maj} (see Figure 1; Koriat, 2012a). This probability can be indexed operationally for each item by (a) the proportion of participants who choose the preferred alternative ("item consensus") or by (b) the proportion of times that the same participant chooses his or her most frequent alternative across several presentations of the item ("item consistency").

Turning next to n_{act}, the number of representations actually drawn, the simulation experiment mentioned earlier indicated that the results for n_{act} mimic very closely those obtained for self-consistency. Assuming that response speed is an inverse function of n_{act}, then response speed should be faster for majority than for minority choices and should vary as a function of p_{maj} and pc_{maj} in much the same way as should confidence judgments (see Koriat, 2012a).

In sum, the basic predictions of SCM are as follows: Confidence and response speed should increase with item consensus—the agreement between participants in making the consensual choice for each item. The same is true for item consistency—the within-person agreement in making the more frequent choice. Item consensus and item consistency are assumed to reflect the polarity of the population of representations associated with each item, and this polarity is assumed to constrain the variability that can be observed in binary decisions for each item. However, when variability in the response choice is observed, confidence and response speed should differ depending on which alternative is chosen: When the decision reached is the decision that accords with that of most other participants, confidence and response speed should be higher than when the decision is a nonconsensual decision. Similarly, in a repeated presentation design, confidence and response speed should be higher for the more frequent response than for the less frequent response. It should be stressed that these predictions are based on the assumption that the same process underlies consensual/frequent decisions and nonconsensual/rare decisions: In each case, each participant chooses the response that is favored by the *majority* of representations in the sample of representations that he/she has retrieved.

Empirical Evidence

In what follows, we present a brief review of the results of several studies that provided a test of the predictions derived from SCM. The aim of some of these studies was to examine the bases of people's subjective confidence and the reasons for their accuracy and inaccuracy. Other studies additionally attempted to use confidence judgments as a tool that could provide insight into the process underlying people's construction of their attitudes, beliefs, preferences, predictions, and category membership decisions (Koriat, 2013; Koriat & Adiv, 2011, 2012; Koriat & Sorka, 2015). We first describe the general methodology used in these studies.

Overview of the Methodology and Analytic Procedure

The procedure in the studies to be reviewed was similar except for the domains of the items used. Participants were presented with a series of 2AFC questions. For each question, they chose one answer and indicated their confidence in their choice either on a full-range scale (0–100) or on a half-range scale (50–100). Response latency was also measured, representing the time it took participants to reach a decision. In all of the studies reviewed in this chapter, participants performed the tasks individually, and had no direct access to the responses of other participants.

The same analytic procedure was applied to the results of all studies (see also Bassili, 2003; Huge & Glynn, 2013). First, the two alternative answers to each item were defined *post hoc* as majority and minority responses on the basis of the distribution of the responses across all participants (items with ties were eliminated). Confidence and response latency were then averaged separately for the majority and minority responses.

All studies provided data regarding the effects of between-individual consensus. In these studies, *item consensus* was defined as the proportion of participants making the majority choice. Item consensus was seen as an index of pc_{maj}. In some studies, the task was repeated several times, between 5 and 7, usually across several sessions that took place on separate days. In these studies, the analyses from the first presentation provided a test of the predictions concerning between-individual consensus, but the analyses across different presentations provided a test of the predictions concerning within-individual consistency. In the latter analyses, the number of times that each of the two responses was made to each item was determined for each participant. The two responses were then classified as *frequent* or *rare* according to their relative frequency across presentations. *Item consistency* was defined as the proportion of times that the frequent choice was made by the person across the repeated presentations of the item, and was used as an alternative index of pc_{maj}.

For some of the tasks used, such as those measuring attitudes and beliefs, the answers do not have a truth-value. These tasks allowed us to test predictions about the *basis* of confidence judgments, but not about their accuracy. Other tasks, for which the answers have a truth-value, provided, in addition, a test of predictions regarding the *accuracy* of confidence judgments. These tasks included word matching, general-information, perceptual comparison, and the prediction of others' responses.

Table 1 The studies reviewed in this chapter. For each study, the table presents an example of an item. It indicates whether the items have a truth-value, and lists the number of items, participants, and presentations used in that study

Study	Example of Item	Truth-Value?	Number of items	Number of Participants	Number of Presentations

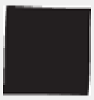
1. Word Matching (Koriat, 1976)	Beautiful Chou Ugly Mei	Yes	85	100	1
2. General Knowledge (Koriat, 2008)	What actress played Dorothy in the original version of the movie The Wizard of Oz? (a) Judy Garland, (b) Greta Garbo	Yes	105	41	1
3. Perceptual—Lines (Koriat, 2011)	Which of the two lines is longer? 	Yes	40	39	5
4. Perceptual—Shapes (Koriat, 2011)	Which of the two geometric shapes has a larger area?  	Yes	40	41	5
5. Predictions of Others' Preferences (Koriat, 2013)	Which sport activity would be preferred by most others? (a) jogging, (b) swimming	Yes	60	41	1
6. Natural Category Membership (Koriat & Sorka, 2015)	Do olives belong to the fruit category?	No	100	33	7
7. Beliefs (Koriat & Adiv, 2012)	There is a supreme being controlling the universe True False	No	60	41	6
8. Attitudes (Koriat & Adiv, 2011)	Capital punishment Yes No	No	50	41	7
9. Personal Preferences (Koriat, 2013)	Which sport activity would you prefer? (a) jogging, (b) swimming	No	60	41	5

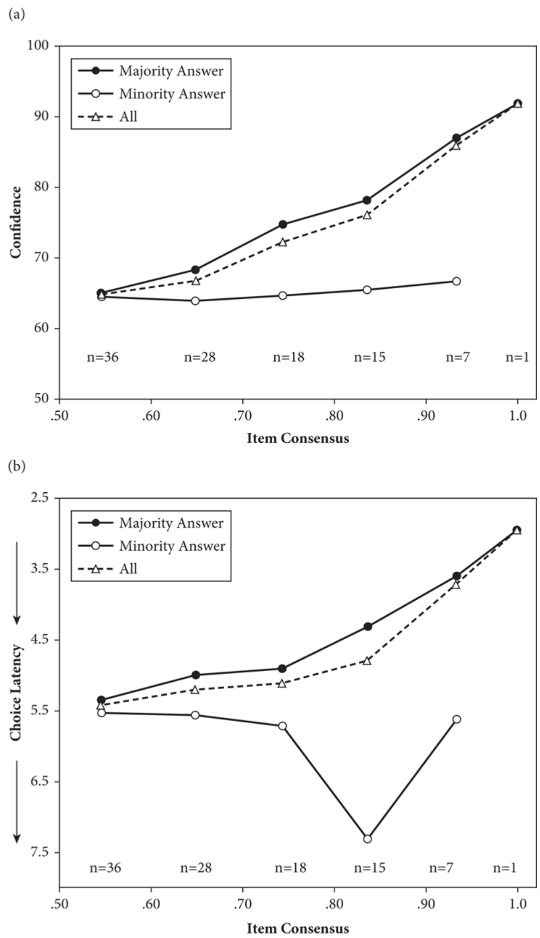
Table 1 lists the studies to be reviewed, and the tasks used in these studies. For each study, it indicates whether the answers have a truth-value, and hence whether the answers could be scored as correct or wrong. The table also indicates the number of items and participants, the confidence scale used, and the number of presentations.

Let us now review the basic findings. We begin with the results for between-person consensus, and then turn to those of within-individual consistency. These results are pertinent to the idea that subjective confidence is based on self-consistency. We then examine the question of the accuracy of subjective confidence. We review the findings regarding the predictions of SCM with regard to metacognitive resolution and metacognitive calibration. SCM will be shown to provide a principled account for observations pertaining to both aspects of the C/A correspondence. We end by examining some general implications of the SCM-based results regarding confidence judgments.

The Relationship of Confidence and Response Latency to Cross-Person Consensus

As noted earlier, pc_{maj} can be indexed by the proportion of participants who choose the majority, consensual answer for each item. To test the predictions of SCM, the following item-based analysis was used. For each item, the answer that was chosen by the majority of participants was designated as the consensual answer, and the other as the nonconsensual answer. Mean confidence was then plotted as a function of item consensus. This was done separately for consensual and nonconsensual answers.

We will illustrate the findings by the results obtained in the study of general-information questions (Koriat, 2008, see Study 2 in Table 1) and then indicate how these findings were replicated for other tasks. In that study, 105 2AFC general-knowledge questions were used. All answers were one or two words long, either a concept or a name of a person or a place. This format was important for the measurement of choice latency (see later). In addition, the questions were chosen deliberately to yield a large number of CW items, for which the wrong answer was likely to be the consensual, majority answer. Confidence was measured.



Click to view larger

Figure 2. Panel A: Mean confidence in the correctness of answers to general-information questions for majority and minority answers and for all responses combined as a function of item consensus (the proportion of participants who chose the majority answer). Panel B presents mean choice latency as a function of item consensus for majority answers, minority answers and for all answers combined. Indicated in the figure is also the number of items (n) in each item consensus category. The results are based on a reanalysis of the data of Koriat (2008). Reproduced with permission from Koriat (2011). Copyright © 2012 by the American Psychological Association.

Figure 2A presents mean confidence judgments for each of six item-consensus categories for both consensual and nonconsensual answers (for one item all participants chose the majority answer). Several trends are suggested by the results:

1. Mean overall confidence judgments ("All" in Figure 2A) increased monotonically with increasing item consensus. When mean confidence and mean item consensus were calculated for each item, the correlation between them over all 105 items was .505, $p < .0001$.
2. However, consensual answers were endorsed with higher confidence (M = 70.9%) than nonconsensual answers (M = 64.6%), $t(103) = 6.74$, $p < .0001$, and this was true regardless of the accuracy of these answers. This difference was consistent across items: For 78 items, confidence was higher for the consensual answer than for the nonconsensual answer compared with 26 items in which the pattern was reversed, $p < .0001$, by a binomial test.
3. It should be noted that in this study, like in all other studies, there were marked and reliable individual differences in the tendency to make relatively high or relatively low confidence judgments (see Kleitman & Stankov, 2001; Stankov & Crawford, 1997). Because the confidence means for consensual and nonconsensual answers in Figure 2A were based on different participants for each item, the differences between these means may reflect a between-individual effect: Participants who tend to choose consensual answers tend to be more confident. To control for inter-participant differences in confidence, the confidence judgments of each participant were standardized so that the mean and standard deviation of each participant were set as those of the raw scores across all participants. Average scores were then calculated for each item for consensual and nonconsensual answers. The consensual-nonconsensual differences were practically the same for the standardized confidence scores.
4. The same general difference between consensual and nonconsensual answers was obtained in subject-based analyses. In these analyses, confidence was compared for each participant between consensual and nonconsensual answers. The results indicated that participants were more confident in their response when that response agreed with the consensual, majority response (72.31%) than when it departed from it (64.36%), $t(40) = 14.79$, $p < .0001$. All 41 participants exhibited this pattern, $p < .0001$, by a binomial test.
5. The moderating effect of item consensus for confidence: We expected the difference in confidence between consensual and nonconsensual responses to increase with item consensus (see Figure 1B). This increase can be seen in 2A but its statistical significance could not be tested on the results presented in that figure because each of the means for the consensual and nonconsensual functions was based on a different combination of participants. However, we calculated for each participant the functions depicted in Figure 2A relating mean confidence in consensual and nonconsensual responses to grouped item consensus categories. The rank order correlation between the ordinal value of the item consensus category (1 to 6) and the difference in mean confidence between consensual and nonconsensual responses (using for each participant the observations for which this difference was computable) averaged .55 across participants, $p < .0001$. This correlation was positive for 35 of the 40 participants (one had a tie), $p < .0001$, by a binomial test.
6. We turn next to the results for response latency. It should be noted that response speed was generally correlated with confidence, consistent with previous findings (e.g., Koriat et al., 2006; Robinson et al., 1997).

Similar analyses to those of confidence were conducted for response latency. The pattern depicted in Figure 2B was largely obtained for response speed. Response speed increased monotonically with item consensus: The correlation between mean latency and item consensus was –.42 across the 105 items, $p < .0001$.

7. Importantly, however, response latencies were longer for nonconsensual answers (5.91s) than for consensual answers (5.14 s), $t(103) = 4.17$, $p < .0001$. This difference was consistent across items: For 67 items, choice latency was longer for nonconsensual answers than for consensual answers compared with 37 items in which the pattern was reversed, $p < .005$, by a binomial test.
8. The same consensual-nonconsensual difference was obtained in subject-based analyses: Participants

responded faster when their answer agreed with the majority answer (5.05s) than when it departed from it (5.79), $t(40) = 6.24$, $p < .0001$. This difference was consistent across participants: For 35 participants, choice latency was shorter for the consensual than for the nonconsensual answer compared with 6 participants for whom the pattern was reversed, $p < .0001$, by a binomial test.

9. The analyses just presented were also repeated after the choice latency scores were standardized. The results yielded essentially the same pattern as that obtained for the raw scores.

10. Like confidence, response latency also yielded a pattern in which the consensual-nonconsensual difference in response latency increased with item consensus. Thus, for consensual answers, the correlation between response latency and item consensus was –.47, $p < .0001$ across the 105 items. The respective correlation for nonconsensual answers was .19, $p < .05$, across 104 items.

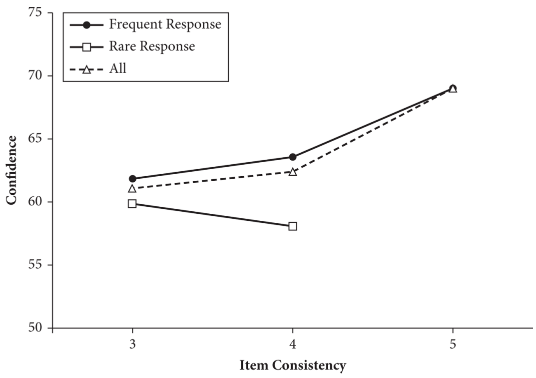
Generality of the results across different tasks.

The pattern of results described for general-information questions was largely replicated across all the tasks listed in Table 1. For each of these tasks, confidence in the consensual response was significantly higher than for the nonconsensual response. This was true in item-based as well as subject-based analyses, and was found even for social attitudes (Koriat & Adiv, 2011) and social beliefs (Koriat & Adiv, 2012) that are known to yield reliable individual differences. Furthermore, in the study of social attitudes, in which the items measured the dimension of Conservatism-Liberalism, the same pattern of results was obtained for the more “liberal” and the more “conservative” participants. In all of the studies listed in Table 1, the consensual-nonconsensual differences were observed even when chronic individual differences were neutralized by standardizing the confidence judgments. For all the tasks, except the personal preferences study (Koriat, 2013), the consensual-nonconsensual difference in confidence increased with item consensus.

Likewise, the results for response latency were also consistent across all of the tasks listed in Table 1. First, consensual responses were associated with significantly faster response times than nonconsensual responses and the difference was obtained even when response latencies were standardized to control for chronic individual differences. For most of the tasks, the consensual-nonconsensual difference increased significantly with item consensus.

The Relationship of Confidence and Response Latency to Within-Person Consistency

According to SCM when participants choose an answer on different occasions, they sample representations from (more or less) the same population of representations on each occasion. Therefore, similar results to those reported previously for consensus should be found when items and answers are categorized in terms of within-person consistency rather than in terms of between-person consensus. The relationship between confidence and within-person consistency will be illustrated by the results obtained in a study using perceptual comparison task (Koriat, 2011, Experiment 1; see Study 3 in Table 1). In that study, 39 participants were presented with pairs of irregular lines and were asked to decide which member of each pair is longer, and to indicate their confidence in the choice. Like the general-information study, the task included several CW-type items, for which most participants chose the wrong answer. The task was administered five times.



Click to view larger

Figure 3. Mean confidence for each participant’s frequent and rare responses for the perceptual judgment task, as a function of item consistency. Reproduced with permission from Koriat (2011). Copyright © 2011

by the American Psychological Association. Reproduced with permission.

To examine the effects of within-person consistency, the number of times that each of the two responses was made to each item was determined for each participant. The two responses were then classified as frequent or rare according to their relative frequency across presentations. In what follows, we summarize the results.

1. Confidence for full consistency responses: All items were classified for each participant into those that elicited the same response across all five blocks (full consistency) and those for which there was some degree of inconsistency (partial consistency). As expected, confidence for the two categories averaged 71.07 and 63.11, respectively, $t(38) = 7.78$, $p < .0001$.
2. Frequent-rare difference in confidence: Figure 3 presents mean confidence for the frequent and rare responses as a function of consistency, that is, the number of times that the frequent response was chosen (3, 4, or 5). Participants were more confident when they chose their more frequent response (62.71) than when they chose their less frequent response (59.03), $t(38) = 5.22$, $p < .0001$. This pattern was exhibited by 31 participants, $p < .0005$, by a binomial test.

3. The moderating effect of item consistency for confidence: The difference in confidence between frequent and rare responses tended to increase with item consistency: Confidence in the frequent response increased significantly with item consistency (3 vs. 4), $t(38) = 2.11$, $p < .05$, whereas confidence in the less frequent response decreased significantly with item consistency, $t(38) = 2.03$, $p < .05$.
4. Essentially, the same pattern was observed for choice latency. Choice latencies were faster for full consistency items (7.14 s) than for partial consistency items (9.25 s), $t(38) = 5.01$, $p < .0001$.
5. The frequent-rare difference in choice latency: Choice latencies were shorter for the frequent responses (5.37 s) than for the rare responses (7.28 s), $t(38) = 5.61$, $p < .0001$. This pattern was exhibited by 37 participants, $p < .0001$, by a binomial test.
6. The moderating effect of item consistency for response latency: Choice latency decreased with item consistency for the frequent choices but increased with item consistency for the rare choices. This pattern roughly mimics the respective pattern for confidence judgments.

Generality of the results across different tasks.

The pattern of results for within-person consistency that was obtained for the perceptual judgments task was largely replicated across all the studies listed in Table 1 in which the task was repeated several times. Confidence judgments were significantly higher for the frequent (majority) response than for the rare (minority) response. In addition, the difference in confidence between frequent and rare responses tended to increase with item consistency for all studies except the preferences study.

The results for response latency were also replicated across most of the tasks listed in Table 1. For all of these studies, frequent responses were generally ventured faster than rare responses. For perceptual judgments, category membership judgments, social beliefs and social attitudes, the frequent-rare difference in choice latency increased significantly with item consistency.

The Postdiction of Confidence and Latency from Response Repetition

It might be argued that the differences observed between frequent and rare choices are due to the changes that occur across blocks: Repeated choices tend to exhibit increased confidence and reduced response latency across repeated presentations (Holland, Verplanken, & van Knippenberg, 2003; Petrocelli, Tormala, & Rucker, 2007). Indeed, in the study of social attitudes (Koriat & Adiv, 2011; see Study 8 in Table 1), confidence tended to increase with repeated presentations. In order to show that the frequent-rare differences in this study were not entirely due to repeated presentations, we attempted to postdict the Block-1 confidence and response latency from the frequency with which the Block-1 choice was made across the subsequent blocks. For each participant, each choice in Block 1 was classified into two categories according to whether it was repeated three times or more in the subsequent six blocks or two times or less. Confidence for the two categories averaged 77.44 and 53.62, respectively, across 39 participants who had both means, $t(38) = 10.15$, $p < .0001$. A similar analysis was carried out for response latency. Response latency in Block 1 averaged 2.80 s for choices that were repeated three times or more, and 4.57 s for those that were repeated two times or less across 36 participants who had both means,

$t(35) = 3.96$, $p < .0005$. Thus, even for Block-1 responses, both confidence and latency discriminated between the more frequent and the less frequent responses: Responses that were made more often across the seven blocks yielded higher confidence and shorter latencies in Block 1 than responses that were made less often.

The Assumption for a Shared Population of Representations

A critical assumption of SCM is that choice and confidence are based on sampling information from a population of item-specific representations that is commonly shared by all participants with the same experience. This assumption is critical for the predictions derived from SCM, and is consistent with similar assumptions made by other researchers (e.g., Gigerenzer et al., 1991). Indeed, previous results suggest that items differ reliably across participants in properties that are relevant to metacognitive judgments (Koriat & Lieblich, 1977). This is the reason why several authors were able to classify items as either “representative” or “deceptive” (Fischhoff et al., 1977; Koriat, 1995, 2008).

Results pertinent to the assumption of a shared population of representations will be illustrated by those obtained in a study of personal preferences (Koriat, 2013; see Study 9 in Table 1). In that study, participants were presented with 2AFC items measuring personal preferences. For each item, they marked their preferred option and indicated their confidence. The task was presented five times.

Interparticipant consensus in choice and confidence.

The assumption that the representations associated with an item are commonly shared implies that properties of items, notably, the likelihood of choosing the majority answer and confidence in that answer, are generally reliable across participants.

Inter-participant reliability for Block 1 was assessed using Cronbach’s alpha coefficient (Crocker & Algina, 1986). This coefficient indicates the degree of agreement among participants in the choice they made for different items and in the confidence associated with their choice. The alpha coefficient was .87 for response choice and .92 for confidence judgments. These high coefficients are in line with the assumption that participants base their choice and confidence on representations that are commonly shared. This was so despite the fact that the task involved personal preferences that tend to be idiosyncratic. The pattern just described for personal preferences was largely replicated across all of the tasks listed in Table 1.

Of course, there might be specific differences between different subgroups of participants that differ in background and degree of expertise. Consider for example the study of social attitudes (Koriat & Adiv, 2011). Because the items in that study were taken from a scale that measures individual differences in conservatism, it was of interest to examine the results for participants who differ in their attitudes. Based on their responses, participants were divided into a group of liberal and a group of conservative respondents. When responses were classified as consensual or nonconsensual on the basis of the responses of all participants, a very similar pattern was observed for the two groups for both confidence and response latency. The results overall were consistent with the idea that regardless of their scoring on the liberal-conservative continuum, participants sampled their clues from a core of representations that is shared by all participants. Clearly, the results may differ for groups that are more clearly divided on the underlying attitude scale. The possibility that groups that clearly differ in their views sample their clues from different populations of clues for each item is worth investigating.

The Relationship Between Consensus and Consistency and Their Joint Effects on Confidence

The relationship between consensus and consistency.

The assumption of a shared population of representations also implies that choices that were made consistently by the same person are also more likely to be made by others. In the personal preferences study (Koriat, 2013), two scores were calculated for each participant for each item: (a) the proportion of times that the choice made in Block 1 was repeated across the subsequent four blocks, and (b) the proportion of other participants (out of 40) who made that choice in Block 1. These two scores were then averaged for each item across participants. The correlation between them (across 59 items) was .34, $p < .01$.

In addition, the confidence of a participant in the choice made in Block 1 predicted the likelihood that that choice would be made by other participants: The correlation was .37, $p < .005$. This correlation suggests that indeed consistency and consensus reflect roughly the same parameter associated with a choice—a parameter that is relevant to confidence in that choice.

Indeed, across several studies, within-person agreement tended to correlate with cross-person consensus: Responses that were made consistently across presentations by one person were also chosen more often by others. This was true for perceptual comparisons, category membership judgments, personal preferences, social attitudes and beliefs. This relationship was taken to support the assumption that within-person agreement and cross-person agreement reflect roughly the same parameter associated with a choice, a parameter that is relevant to confidence in that choice.

Consensus versus consistency.

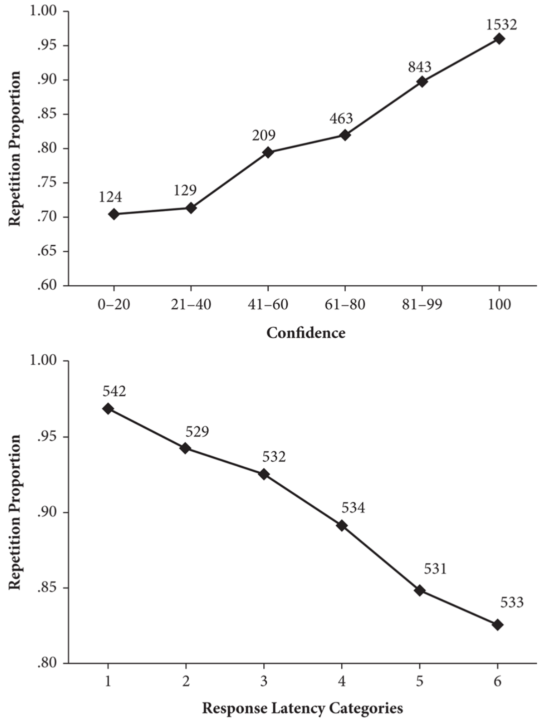
The results just described suggest that both item consensus and item consistency can be taken to reflect pc_{maj} and hence are equally diagnostic of p_{maj}—the polarization of the population of representations associated with each item.

However, in certain domains, consistent individual differences exist. In such domains, within-person agreement for each item might prove to be a better diagnostic of the self-consistency underlying choice and confidence than is cross-person agreement. Indeed, such was suggested to be the case for social attitudes, social beliefs, and personal preferences (Koriat, 2013; Koriat & Adiv, 2011, 2012). The results obtained for these domains indicated that confidence is predicted better from within-person agreement than from cross-person agreement. To illustrate, for the preferences study, response consistency had a much stronger effect than response consensus: For the effect of consistency, the partial $\eta 2$, as an estimate of effect size, was .89, whereas that for consensus was only .14.

Confidence and Response Latency as Predictors of Reproducibility

The assumption underlying SCM is that, like statistical level of confidence, subjective confidence represents a subjective assessment of reproducibility—the likelihood that a new sample of representations drawn from the same population will yield the same choice. It was proposed that although confidence judgments are construed subjectively as pertaining to validity—the likelihood that the answer is correct, they actually monitor reliability or reproducibility.

The idea that confidence and choice latency are predictors of reproducibility will be illustrated by the results obtained in the study involved category membership judgments (Koriat & Sorka, 2015; see Study 6 in Table 1). Participants were presented with an item (e.g., *Avocado*). They were asked to judge whether it was a member of a given natural category (e.g., FRUIT) or not, and to indicate their confidence. The confidence judgments in Block 1 were grouped into six categories, and the proportion of response repetitions—the likelihood of making the same response over the subsequent six blocks—was calculated across all participants. The results, pooled across participants and items, are presented in Figure 4A. Indicated in this figure is also the number of observations in each category. The function indicates that response repetition increased monotonically with confidence in Block 1. The Spearman rank-order correlation over the six values was 1.0, $p < .0001$. When Pearson correlation was calculated for each participant across the full range of confidence judgments, the correlation averaged .37, $p < .0001$. This correlation was positive for all 33 participants, $p < .0001$, by a binomial test.



Click to view larger

Figure 4. Panel A presents the likelihood of repeating the Block-1 choice across the subsequent six blocks (repetition proportion) for each of six confidence categories for the category membership judgments experiment. Panel B plots repetition proportion as a function of Block-1 response latency. Indicated in the body of this figure is also the number of observations in each category. Reproduced with permission from Koriat and Sorka (2015). Copyright © 2014 by Elsevier.

A similar analysis was carried out for response latency. The response latencies in Block 1 were grouped as before. Figure 4B presents mean repetition proportion as a function of response latency. The Spearman rank-order correlation across the six points was –1.0, $p < .0001$. The Pearson correlation, calculated for each participant across the full range of response latencies averaged –.25, $p < .0001$. This correlation was positive for 32 participants out of the 33 participants, $p < .0001$, by a binomial test. Similar results were obtained in all the studies in which the task was administered several times (see Koriat, 2011, figure 3; Koriat, 2012a, 2012 figures 9, 10; Koriat & Adiv, 2011, figure 7; Koriat & Adiv, 2012, figure 6).

The Accuracy of Confidence Judgments: The Consensuality Principle

Examination of the relationship between confidence and accuracy for tasks in which the answer has a truth-value clearly supported the consensuality principle: For CC items, in which the correct response was the consensual response, confidence was consistently higher for correct responses than for wrong responses. For CW items, in contrast, for which the wrong response was the consensual response, confidence was higher for the wrong responses. Thus, confidence is correlated with the consensuality of the answer rather than with its accuracy. The consensuality principle was demonstrated for general-information questions (Koriat, 2008), FOK judgments (Koriat, 1995), and perceptual judgments (Koriat, 2011). It was also observed for sentence memory (Brewer & Sampaio, 2006) and for the prediction of others' personal preferences (Koriat, 2013). Both confidence and response speed were correlated with the consensuality of the choice rather than with its correctness.

The consensuality principle can be illustrated by the results for the perceptual comparison task (Koriat, 2011, Experiment 1; see Study 3 in Table 1). Confidence judgments for correct and wrong responses were averaged for each participant for the CC and CW items. The results yielded a crossover interaction. For the CC items, confidence was higher for correct answers (69.09) than for wrong answers (60.58), $t(30) = 5.13$, $p < .0001$, as would be expected. In contrast, for the CW items, confidence was significantly higher for the wrong answers (65.29) than for the correct answers (58.81), $t(30) = 2.68$, $p < .05$. We also calculated the within person C/A gamma correlation across the 40 items. Mean gamma correlation was positive across the CC items, (.32), but it was significantly negative (–.25) across the CW items.

These results suggest that the positive C/A correlation that has been observed in numerous studies is due to the fact that in these studies participants' judgments are generally correct. Such is generally the case for psychophysical tasks and for general-information questions. For these tasks, the positive relationship observed between confidence and accuracy is possibly mediated by the relationship of confidence to self-consistency.

The Calibration of Confidence Judgments

The observation that has attracted much in the study of the calibration of confidence judgments is the overconfidence bias: Confidence judgments tend to be inflated relative to accuracy. Among the explanations that were proposed for this bias is that it derives from overreliance on the strength rather than the weight of evidence (Griffin & Tversky, 1992), that it reflects self-serving motivations (Taylor & Brown, 1988; see Metcalfe, 1998; Nickerson, 1998) and that it is due to random noise in participants' judgments (Erev, Wallsten, & Budescu, 1994; Soll, 1996). Some researchers also argued that the overconfidence bias actually represents a methodological artifact that stems from researchers' tendency to include tricky or deceptive items in the experimental sample (Björkman, 1994; Gigerenzer et al., 1991; Juslin, 1993, 1994).

SCM offers a simple account of the overconfidence bias. This bias is assumed to stem from the reliability-validity discrepancy: Confidence monitors reliability (or self-consistency) but its accuracy is evaluated in calibration studies against correctness. As stated in many textbooks, reliability sets an upper limit on validity so that reliability is practically always higher than validity. Indeed, the evaluation of confidence against indexes of self-consistency yielded a markedly smaller tendency toward overconfidence than when confidence was compared to accuracy (see Figure 13, Koriat, 2012a,).

Some General Implications

The results reviewed in this chapter have several general implications that will be mentioned.

The Online Construction of Judgments

In some of the studies listed in Table 1, confidence judgments were used primarily to provide some clues regarding the processes underlying choice and decisions. Consider, for example, the study to social attitudes (Koriat & Adiv, 2011).

The traditional view to social attitudes has treated attitudes as evaluative predispositions that are relatively stable over time (Allport, 1935). In contrast, the approach that has been gaining in popularity views attitudes as judgments that are formed on the spot; therefore, they may vary depending on the person's current goals, mood, and context (Bless, Mackie, & Schwarz, 1992; Schwarz, 1999, 2007; Schwarz & Bohner, 2001; Schwarz & Strack, 1985; Wilson & Hodges, 1992). The attitudinal construal view has led to increased emphasis on the malleability and context-sensitivity of attitudinal judgments. A similar view has been advanced with regard to personal preferences: Preferences are constructed in the process of elicitation rather than retrieved ready-made from memory (Lichtenstein & Slovic, 2006; Slovic, 1995).

Assuming that attitudinal judgments are constructed on the fly on the basis of the associations that come to mind at the time of making a judgment, some fluctuation in attitudinal judgments may be expected even in the absence of any manipulation that attempts to affect these judgments. The results for the social attitudes study (Koriat & Adiv, 2011) indicated that attitude certainty and response latency are diagnostic of both the stable and variable aspects of attitudinal judgments. The stable aspects are reflected in the systematic functions relating mean self-consistency and mean n_{act} to p_{maj}. The variable aspects, which stem from sampling fluctuations, are disclosed by the systematic differences between majority and minority choices in both self-consistency and n_{act}. Thus, SCM can provide some clues to the construction of attitudinal judgment.

These assumptions have been extended to the study of social beliefs (Koriat & Adiv, 2012) and category membership decisions (Koriat & Sorka, 2015). The assumption is that in both cases, participants construct their judgments on the spot depending on the clues that are accessible in making a judgment. In both studies, confidence and response latency were found to track the stable and variable contributions to judgments.

Group Decisions

Studies that compared individual and group decisions have yielded somewhat inconsistent results. The *groupthink phenomenon*, which refers to a mode of decision-making that occurs within a cohesive group, has been claimed to underlie some of the disastrous decisions made in US history (Baron, 2005; Esser, 1998; Janis, 1982). However, several studies have indicated that cooperative groups perform better than independent individuals on a wide range of problem-solving tasks (e.g., Hill, 1982; Laughlin, Hatch, Silver, & Boh, 2006). In particular, Bahrami et al. (2010) compared individual and dyadic performance in a simple visual task. Participants performed the task in dyads. They first made their decision individually, then shared their decisions, and reached a joint decision. The results indicated that “two heads were definitely better than one provided they were given the opportunity to communicate freely.”

SCM was shown to predict the dyadic superiority effect in the absence of any communication between the members of a dyad, assuming that for each item, the dyadic decision is dominated by the individual with higher confidence (see Koriat, 2012c). A maximum-confidence slating (MCS) algorithm was used for combining judgments across two people who operate individually. For each trial, the decision that was made with higher confidence by one member of the dyad was selected, circumventing dyadic interaction altogether. Indeed, for a representative set of items, as well as for sets of CC items, performance based on the MCS heuristic was better than that of the best performing member of a dyad. However, for CW sets of items, the MCS algorithm yielded *worse* performance even than the worst of the two members of a dyad. These results follow from the consensuality principle: For CC items high-confidence responses are generally correct whereas for CW items high-confidence responses are generally wrong.

The Effects of Group Pressure toward Conformity

Extensive research in social psychology has demonstrated dramatic effects of group consensus on the judgments of individual members. In particular, results suggest a prototypical majority effect (PME) for confidence judgments and response speed. First, majority views are endorsed with greater confidence and are expressed with greater fluency than nonconsensual, minority views. Second, the difference between majority and minority responses in both confidence and response speed increases with the size of the majority. Bassili (2003) observed a “minority slowness effect”: People who hold a minority opinion tend to express that opinion less quickly than those who hold the majority opinion. He attributed this effect to social inhibition deriving from conformity pressures. This interpretation of PME implies a causal relationship: Group unanimity influences the confidence of individuals in their own views and the ease with which they express these views (see also Huge & Glynn, 2013).

However, the results reviewed in this chapter (see Koriat, Adiv, & Schwarz, under revision) suggest that the PME for 2AFC items can arise from the very process underlying subjective confidence and response speed independent of any social influence. Indeed, as reviewed in this chapter, PME was demonstrated for tasks and conditions that are stripped of social relevance. It was also observed in within-individual analyses when majority and minority responses were defined in terms of the relative distribution of these responses across repeated presentations. It was proposed that the majority effects that have been observed in previous studies might include two components. The first, a process-based, internally driven PME component, which derives from the process underlying confidence in 2AFC problems. The second, externally driven PME component is due specifically to social influence. Whereas the former component is expected to transpire generally for all tasks alike, the latter component should be particularly pronounced for tasks for which the response is prone to the effects of social influence.

Philosophical Implications

The question of certainty in knowledge has received a great deal of attention in the philosophy of knowledge in the context of discussions about truth and its justification. The work reviewed in this chapter may also have some philosophical implications (see Koriat & Adiv, 2012).

A central issue in the philosophy of knowledge is associated with the traditional distinction between rationalism and empiricism (see Edwards, 1996; Markie, 2008). The rationalist approach focuses on intuitive knowledge—a priori propositions whose truth is self-evident, independent of sense experience. In contrast, empiricists argued that the origin of knowledge resides in the external world. According to them, sense experience is the ultimate source of knowledge and therefore the focus should be on a posteriori propositions whose justification relies on empirical observations. SCM suggests that although the validation of one's own knowledge is based on retrieving information from memory, the underlying process is actually analogous to that in which information is sampled from the outside world with the goals of (a) testing a hypothesis about the population from which the sample was drawn and (b) assessing the likelihood that the conclusion reached is correct. Such is the case whether participants need to validate propositions whose truth is a priori, such as logical assertions or metaphysical beliefs, or propositions whose truth is a posteriori, dependent on experience. Thus, the prototype for the underlying process is provided by the statistical procedures that are used by researchers in attempting to draw conclusions about the external world: A proximal sample of observations is used to make inferences about some "true" parameter of a distal population. The critical difference, of course, is that information is sampled from within rather than from without.

Another distinction in the philosophy of knowledge is the distinction between two philosophical theories of truth, correspondence theories and coherence theories (Kirkham, 1992). Correspondence theories posit that the truth or falsity of a statement is determined only by how the statement relates to the world, and whether it accurately describes objects or facts. Coherence theories, in contrast, assume that the truth or falsity of a statement is determined by its relations to other statements rather than its relation to the world. In this view, a person's belief is true if it is coherent with his or her body of beliefs, that is, if it is a constituent of a systematically coherent whole. SCM implies a specific relationship between coherence and correspondence: Although subjective confidence in the truth of a statement pertains to correspondence (e.g., that "Sydney is indeed the capital of Australia"), the mnemonic cue for metacognitive assessments of correspondence is degree of coherence. Confidence in a belief or answer depends on the extent to which the various pieces of information that come to mind fit together with that belief or answer as well as with one another. Because people have no access to the object of their beliefs over and above what they know about it, they rely on a fast assessment of overall coherence (see Bolte & Goschke, 2005) as a basis for their judgments about correspondence.

In sum, SCM was initially motivated by the attempt to clarify the processes underlying warranted and unwarranted convictions. However, it was extended to examination of the basis of confidence judgments in tasks for which the response does not have a truth-value. The results so far have been very consistent across a variety of tasks, supporting the generality of the theoretical framework underlying SCM.

Future Work

Although the SCM model is very rudimentary, and incorporates strong assumptions, it yielded a large number of gross predictions that were generally supported across several tasks. Future work, however, must attempt to refine the model in order to allow more detailed, quantitative predictions. It is also of interest to see to what extent the model can be extended to tasks other than 2AFC tasks. For example, it is important to examine whether it can be applied to test formats involving more than two alternative options.

Authors' Note

The preparation of this chapter was supported by a grant (number 2013039) from the United States—Israel Binational Science Foundation to Asher Koriat and Norbert Schwarz. We are grateful to Miriam Gil for her help in the analyses, and to Etti Levran (Merkine) and Ornit Tzuri for their help in copyediting.

Correspondence concerning this article should be addressed to Asher Koriat, Department of Psychology, University of Haifa, Haifa 3498838, Israel. E-mail: akoriat@research.haifa.ac.il

References

Alba, J. W., & Marmorstein, H. (1987). The effects of frequency knowledge on consumer decision making. *Journal of Consumer Research*, *14*, 14–25.

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *A handbook of social psychology* (pp. 789–844). Worcester, MA: Clark University Press.

Audley, R. J. (1960). A stochastic model for individual choice behavior. *Psychological Review*, *67*, 1–15.

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting mind. *Science*, *27*, 1022–1023.

Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 929–945.

Baron, R. S. (2005). So right it's wrong: Groupthink and the ubiquitous nature of polarized group decision making. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 219–253). San Diego, CA: Elsevier Academic Press.

Bassili, J. N. (2003). The minority slowness effect: Subtle inhibitions in the expression of views not shared by others. *Journal of Personality and Social Psychology*, *84*, 261–276.

Benjamin, A. S., & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 309–338). Mahwah, NJ: Erlbaum.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*, 55–68.

Björkman, M. (1994). Internal cue theory: Calibration and resolution of confidence in general knowledge. *Organizational Behavior and Human Decision Processes*, *58*, 386–405.

Bless, H., Mackie, D. M., & Schwarz, N. (1992). Mood effects on encoding and judgmental processes in persuasion. *Journal of Personality and Social Psychology*, *63*, 585–595.

Bolte, A., & Goschke, T. (2005). On the speed of intuition: Intuitive judgments of semantic coherence under different response deadlines. *Memory & Cognition*, *33*, 1248–1255.

BonJour, L. (1985). *The structure of empirical knowledge.* Cambridge, MA: Harvard University Press.

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, *72*, 691–695.

Brewer, N., Keast, A., & Rishworth, A. (2002). The confidence-accuracy relationship in eyewitness identification: The effects of reflection and disconfirmation on correlation and calibration. *Journal of Experimental Psychology: Applied*, *8*, 46–58.

Brewer, W. F., & Sampaio, C. (2006). Processes leading to confidence and accuracy in sentence recognition: A metamemory approach. *Memory*, *14*, 540–552.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, *7*, 26–48.

Chandler, C. C. (1994). Studying related pictures can reduce accuracy, but increase confidence, in a modified recognition test. *Memory & Cognition*, *22*, 273–280.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York, NY: Holt, Rinehart, & Winston.

Dhami, M., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130*, 959–988.

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*, 69–106.

Edwards, P. (Ed.) (1996). *The encyclopedia of philosophy*, Vols. 1 and 2—Complete and unabridged. New York, NY & London, England: Macmillan Reference.

Engel, P. (1998). Believing, holding true, and accepting. *Philosophical Explorations*, *2*, 140–151.

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, *101*, 519–527.

Esser, J. K. (1998). Alive and well after 25 years: A review of groupthink research. *Organizational Behavior and Human Decision Processes*, *73*, 116–141.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 552–564.

Fisher, R. A. (1925). *Statistical Methods for Research Workers.* London, England: Oliver & Boyd.

Galton, F. (1907). Vox populi. *Nature*, *75*, 450–451.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*, 20–29.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506–528.

Griffin, D., & Brenner, L. (2004). Perspectives on probability judgment calibration. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 177–198). Malden, MA: Blackwell.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*, 411–435.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, 208–216.

Hill, G. W. (1982). Group versus individual performance: Are N+1 heads better than one? *Psychological Bulletin*, *91*, 517–539.

Hoffrage, U. (2004). Overconfidence. In R. F. Pohl (Ed.), *Cognitive illusions: Fallacies and biases in thinking, judgment, and memory* (pp. 235–254). Hove, England: Psychology Press.

Holland, R. W., Verplanken, B., & van Knippenberg, A. (2003). From repetition to conviction: Attitude accessibility as a determinant of attitude certainty. *Journal of Experimental Social Psychology*, *39*, 594–601.

Huge, M. E., & Glynn, C. J. (2013). Hesitation blues: Does minority opinion status lead to delayed responses? *Communication Research*, *40*, 287–307.

Janis, I. (1982). *Groupthink: A psychological study of policy decisions and fiascos*, 2nd ed. Boston, MA: Houghton Mifflin.

Juslin, P. (1993). An explanation of the hard-easy effect in studies of realism of confidence in one's general knowledge. *European Journal of Cognitive Psychology*, *5*, 55–71.

Juslin, P. (1994). The overconfidence phenomenon as a consequence of informal experimenter-guided selection of almanac items. *Organizational Behavior and Human Decision Processes*, *57*, 226–246.

Juslin, P., & Olsson, H. (1997). Thurstonian and Brunswikian origins of uncertainty in judgment: A sampling model of confidence in sensory discrimination. *Psychological Review*, *104*, 344–366.

Juslin, P., Winman, A., & Olsson, H. (2003). Calibration, additivity, and source independence of probability judgments in general knowledge and sensory discrimination tasks. *Organizational Behavior and Human Decision Processes*, *92*, 34–51.

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*, *35*, 157–175.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*, 1–24.

Kleitman, S., & Stankov, L. (2001). Ecological and person-oriented aspects of metacognitive processes in test-taking. *Applied Cognitive Psychology*, *15*, 321–341.

Keren, G. (1991). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica*, *77*, 217–273.

Kirkham, R. L. (1992). *Theories of truth.* Cambridge, MA: MIT.

Koriat, A. (1975). Phonetic symbolism and the feeling of knowing. *Memory & Cognition*, *3*, 545–548.

Koriat, A. (1976). Another look at the relationship between phonetic symbolism and the feeling of knowing. *Memory & Cognition*, *4*, 244–248.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*, 311–333.

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

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, *9*, 149–171.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). New York, NY: Cambridge University Press.

Koriat, A. (2008). Subjective confidence in one's answers: The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 945–959.

Koriat, A. (2011). Subjective confidence in perceptual judgments: A test of the self-consistency model. *Journal of Experimental Psychology: General*, *140*, 117–139.

Koriat, A. (2012a). The self-consistency model of subjective confidence. *Psychological Review*, *119*, 80–113.

Koriat, A. (2012b). The subjective confidence in one's knowledge and judgments: Some metatheoretical considerations. In M. J. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *The foundations of metacognition* (pp. 213–233). Oxford, England: Oxford University Press.

Koriat, A. (2012c). When are two heads better than one and why? *Science*, *336*, 360–362.

Koriat, A. (2013). Confidence in personal preferences. *Journal of Behavioral Decision Making*, *26*, 247–259.

Koriat, A., & Adiv, S. (2011). The construction of attitudinal judgments: Evidence from attitude certainty and response latency. *Social Cognition*, *29*, 577–611.

Koriat, A., & Adiv, S. (2012). Confidence in one's social beliefs: Implications for belief justification. *Consciousness and Cognition*, *21*, 1599–1616.

Koriat, A., Adiv, S., & Schwarz, N. (2015). Views that are shared with others are expressed with greater confidence and greater fluency independent of any social influence (under revision).

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 483–502). New York, NY: Guilford Press.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology:*

Human Learning and Memory, *6*, 107–118.

Koriat, A., & Lieblich, I. (1977). A study of memory pointers. *Acta Psychologica*, *41*, 151–164.

Koriat, A., Ma’ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacogntion: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, *135*, 36–69.

Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments: Evidence from subjective confidence. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of memory and metamemory* (pp. 117–135). New York, NY: Psychology Press.

Koriat, A., Pansky, A., & Goldsmith, M. (2011). An output-bound perspective on false memories: The case of the Deese–Roediger–McDermott (DRM) paradigm. In A. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork* (pp. 301–332). London, England: Psychology Press.

Koriat, A., & Sorka, H. (2015). The construction of categorization judgments: Using subjective confidence and response latency to test a distributed model. *Cognition*, *134*, 21–38.

Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems: Effects of group size. *Journal of Personality and Social Psychology*, *90*, 644–651.

Liberman, V., & Tversky, A. (1993). On the evaluation of probability judgments: Calibration, resolution, and monotonicity. *Psychological Bulletin*, *114*, 162–173.

Lichtenstein, S., & Slovic, P. (2006). *The construction of preference.* London, England: Cambridge University Press.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York, NY: Cambridge University Press.

Lindsay, R. C., Wells, G. L., & Rumpel, C. M. (1981). Can people detect eyewitness-identification accuracy within and across situations? *Journal of Applied Psychology*, *66*, 79–89.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, *70*, 151–159.

Markie, P. (2008). Rationalism vs. empiricism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Palo Alto, CA: The Center for the Study of Language and Information (CSLI), Stanford University. **http://plato.stanford.edu/archives/fall2008/entries/rationalism-empiricism/**

McKenzie, C. R. M. (1997). Underweighting alternatives and overconfidence. *Organizational Behavior and Human Decision Processes*, *71*, 141–160.

McKenzie, C. R. M. (2005). Judgment and decision making. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of cognition* (pp. 321–338). London, England: Sage.

Metcalfe, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics? *Personality and Social Psychology Review*, *2*, 100–110.

Metcalfe, J. (2000). Feelings and judgments of knowing: Is there a special noetic state? *Consciousness and Cognition*, *9*, 178–186.

Metcalfe, J., & Dunlosky, J. (2008). Metamemory. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory. Vol. 2 of Learning and memory: A comprehensive reference*, 4 vols. (J. Byrne, Editor) (pp. 349–362). Oxford, England: Elsevier.

Mickes, L., Hwe, V., Wais, P. E., & Wixted, J. T. (2011). Strong memories are hard to scale. *Journal of Experimental Psychology: General*, *140*, 239–257.

Mozer, M. C., Pashler, H., & Homaei, H. (2008). Optimal predictions in everyday cognition: The wisdom of individuals or crowds? *Cognitive Science*, *32*, 1133–1147.

Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology*, *12*, 595–600.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133.

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*, 267–270.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, *68*, 29–46.

Petrocelli, J. V., Tormala, Z. L., & Rucker, D. D. (2007). Unpacking attitude certainty: Attitude clarity and attitude correctness. *Journal of Personality and Social Psychology*, *92*, 30–41.

Read, J. D., Lindsay, D. S., & Nicholls, T. (1998). The relation between confidence and accuracy in eyewitness identification studies: Is the conclusion changing? In C. P. Thompson, D. J. Herrmann, J. D. Read, D. Bruce, D. G. Payne, & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives* (pp. 107–130). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

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

Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. Journal of Applied Psychology, *82*, 416–425.

Ross, M. (1997). Validating memories. In N. L. Stein, P. A. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp. 49–81). Mahwah, NJ: Erlbaum.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93–105.

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, *25*, 638–656.

Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser, & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (Vol. 1, pp. 436–457). Oxford, England: Blackwell.

Schwarz, N., & Strack, F. (1985). Cognitive and affective processes in judgments of well-being: A preliminary model. In H. Brandstatter & E. Kirchler (Eds.), *Economic psychology* (pp. 439–447). Linz, Austria: Taubler.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11–36.

Slobin, D. (1968). Antonymic phonetic symbolism in three natural languages. *Journal of Personality and Social Psychology*, *10*, 301–305

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*, 364–371.

Soll, J. B. (1996). Determinants of overconfidence and miscalibration: The roles of random error and ecological structure. *Organizational Behavior and Human Decision Processes*, *65*, 117–137.

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*, 315–327.

Stankov, L., & Crawford, J. D. (1997). Self-confidence and performance on tests of cognitive abilities. *Intelligence*, *25*, 93–109.

Stewart, N. (2009). Decision by sampling: The role of the decision environment in risky choice. *Quarterly Journal of Experimental Psychology*, *62*, 1041–1062.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, *53*, 1–26.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Tormala, Z. L., & Rucker, D. D. (2007). Attitude certainty: A review of past findings and emerging perspectives. *Social and Personality Psychology Compass*, *1*, 469–492.

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, *21*, 437–484.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness & Cognition*, *18*, 22–38.

Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 582–600.

Vickers, D. (2001). Where does the balance of evidence lie with respect to confidence? In E. Sommerfeld, R. Kompass, & T. Lachmann (Eds.), *Proceedings of the seventeenth annual meeting of the international society for psychophysics* (pp. 148–153). Lengerich: Pabst Science Publishers.

Vickers, D., & Pietsch, A. (2001). Decision making and memory: A critique of Juslin and Olsson's (1997) sampling model of sensory discrimination. *Psychological Review*, *108*, 789–804.

Vickers, D., Smith, P., Burt, J., & Brown, M. (1985). Experimental paradigms emphasising state or process limitations: II effects on confidence. *Acta Psychologica*, *59*, 163–193.

Vul, E., Goodman, N. D., Griffiths, T. L., & Tenenbaum, J. B. (2009). One and done? Optimal decisions from very few samples. *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 148–153). Austin, TX: Cognitive Science Society.

Wallsten, T. S., Budescu, D. V., Erev, I., & Diederich, A. (1997). Evaluating and combining subjective probability estimates. Journal of Behavioral Decision Making, *10*, 243–268.

Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 37–65). Hillsdale, NJ: Erlbaum.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*, 1025.

Yzerbyt, V. Y., Lories, G., & Dardenne, B. (Eds.). (1998). *Metacognition: Cognitive and social dimensions.* London, England: Sage.

Asher Koriat

Asher Koriat, Department of Psychology, University of Haifa

Shiri Adiv

Shiri Adiv, Department of Psychology, University of Haifa