TARGET ARTICLE
Precognition as a Form of Prospection: A Review of the Evidence

Julia A. Mossbridge
Northwestern University and Institute of Noetic Sciences, Petaluma, California

Dean Radin
Institute of Noetic Sciences, Petaluma, California

Prospection, the act of attempting to foresee one’s future, is generally assumed to be based on conscious and nonconscious inferences from past experiences and anticipation of future possibilities. Most scientists consider the idea that prospection may also involve influences from the future to be flatly impossible due to violation of common sense or constraints based on one or more physical laws. We present several classes of empirical evidence challenging this common assumption. If this line of evidence can be successfully and independently replicated using preregistered designs and analyses, then the consequences for the interpretation of experimental results from any empirical domain would be profound.

Keywords: prospection, precognition, presentiment, anticipatory activity, retrocausality

In this review, we discuss a set of controlled experiments investigating what we will argue is an inherent human ability that allows for accurate prediction of future events without inferential means; in the vernacular this ability is known as precognition. While taking this line of research seriously may seem beyond the pale to some, it is worth remembering that advances in psychology and physics have repeatedly demonstrated that everyday intuitions about the nature of reality only partially reflect the nature of reality itself. It is possible that such imprecise intuitions include the concept of a fixed “arrow of time,” which Einstein famously called a “stubbornly persistent illusion” (Einstein & Hawking, 2007).

Everyday Intuitions About Events in Time

Common sense intuitions about events unfolding in time include the following:

- Events in the physical world are mirrored by our almost immediate perceptual recreation of those events in essentially the same order that they occurred.
- The recreation of physical events, first in perception and later in memory, occurs in a linear order based upon our original perception of events. Thus, given two events and assuming perfect perception and memory, Event A is said to occur before Event B if at some point we remember Event A but we have not yet experienced Event B, and then later, after Event B occurs, we remember both events.
- Event A may only be said to “cause” Event B if Event A precedes Event B.
- What we remember has occurred in the past.

These everyday intuitions seem reasonable because they arise and are inculcated by innumerable conscious waking experiences (Mossbridge, 2015). However, as developments in psychology and physics have repeatedly shown, even reasonable assumptions should be thoroughly checked using multiple methods to determine if in fact they are correct.

What methods can allow us to double-check our intuitions about the nature of time? If our everyday intuitions are correct and we assume that certain future events cannot be inferred from present circumstances or extrapolations

Julia A. Mossbridge, Department of Psychology, Northwestern University, and Institute of Noetic Sciences, Petaluma, California; Dean Radin, Institute of Noetic Sciences.

The first author was supported by Bial fellowship 97/16 and the second author by Bial fellowship 260/14. We thank Caroline Watt and Imants Barušs for providing useful comments on portions of this article.

Correspondence concerning this article should be addressed to Julia A. Mossbridge, Department of Psychology, Northwestern University, Swift Hall, Room 102, 2029 N Campus Drive, Evanston, IL 60208. E-mail: jmossbridge@gmail.com
based on past circumstances, and if we also recognize that coincidences will occasionally occur by chance, then the following predictions should be borne out in carefully conducted experiments: (a) Dreams will bear no relationship to unpredictable future events beyond chance levels; (b) individuals cannot consciously out-guess unpredictable future events at a rate greater than chance; (c) behavioral measures administered in the present will not be affected by future events; (d) physiological measures will not be affected by future events; (e) events in the future do not influence what occurs in the present, except in cases of prospective planning (e.g., “I need to prepare for rain tomorrow, so today I will purchase an umbrella”).

Precognition is not required to guess that researchers have performed experiments that examine all of these predictions, nor that some of these studies suggest that widely held intuitions about time are empirically falsifiable. Here we will describe and evaluate the literature in domains known as precognitive dreaming, forced-choice precognition, free-response precognition, implicit precognition, and presentiment. We will build from what we believe to be the weakest to the strongest evidence for precognition. Along the way, and in a concentrated form at the end of this review, we will suggest methodological improvements and future directions, as well as present some speculations about potential mechanisms.

**Tests of Precognitive Dreaming**

Anecdotal claims of precognitive dreams are common, reported by ~17–58% of survey participants (Lange, Schredl, & Houran, 2000; Parra, 2013). Many such claims have been reported, including dreams of historical figures such as Abraham Lincoln and Mark Twain. In the former case, on each of 3 nights before he was assassinated, the bodyguard assigned to Abraham Lincoln reported that the president mentioned dreams of his death (Lewis, 1994). In the latter case, Mark Twain wrote that he dreamed of his brother’s death before his brother was tragically killed in a steamboat accident (Zohar, 1982). These stories are intriguing, but because of known frailties of memory, including confabulation and elaboration, and because some events are predictable based on nonconscious associations drawn from sensory cues (e.g., Bechara, Tranel, Damasio, & Damasio, 1996), anecdotes alone cannot provide persuasive evidence that precognition actually occurs in dreams.

For scientifically valid evidence, we turn to controlled experiments. Researchers have published four well-controlled precognitive dream experiments in peer-reviewed journals. By “well-controlled,” we mean the following: (a) Participants were asked to attempt to dream about a target they would see the next day (i.e., images, staged multisensory experiences, or video clips); (b) a random-number generator was used to select one target from a pool of at least four available targets (e.g., a video clip of a birthday party) and only that one target was later shown to the dreamer; (c) on each trial the experimenters selected the target only after dream reporting was complete and submitted to the experimenters, and before the experimenters read the dream reports; (d) independent judges naïve to the identity of the actual target judged the similarity between dream content and the target; (e) judges’ responses were considered final. Researchers reported significant results using binomial statistics (α = 0.05, two-tailed tests) in three out of four of these studies (Krippner, Honorton, & Ullman, 1972; Krippner, Ullman, & Honorton, 1971; Watt, 2014). The fourth study (Watt, Wiseman, & Vuillaume, 2015) did not show statistically significant results but the effect was in the predicted direction (ES = 0.11; N = 20 with one trial per person).

Four other peer-reviewed experiments deserve consideration (Luke, 2002; Luke & Zychowicz, 2014; Luke, Zychowicz, Richterova, Tjurina, & Polonnikova, 2012; Sherwood, Roe, Simmons, & Biles, 2002), although they do not fit our methodological constraints. The difference is that in these additional studies, dreamers viewed all items in the target pool after their dreams had been recorded. This allowed each dreamer to rank the similarity of all of the items in the target pool against his or her own dream content, which in turn simplified the judging of
dream accuracy. But it also violated requirement (a) above. This type of design may pose a problem if dreams can indeed gather information about a future event. This is because all potential targets, rather than just the actual target, might confuse underlying precognition mechanisms in that they might send information about multiple targets backward in time. This postulated “information confusion” could explain why none of these studies produced a significant result. It is worth reminding readers that the investigator’s goal when exploring a hypothesis is to try to disprove that hypothesis. If the hypothesis is that people can precognize future events, then it is not useful for researchers to expose participants to several future events, because according to the hypothesis the researcher is trying to disprove, participants should be able to precognize any or all of those events.

Further description of Watt’s (2014) study can illustrate the important elements of a well-controlled precognitive dream experiment, as well as a methodological concern that was addressed by the authors in a timely fashion. The targets in this study consisted of a group of 68 video clips the experimenters developed before they reviewed participants’ dreams. The study was preplanned to include 200 trials, with four trials performed by each of the first 50 participants. Möck (2015) raised the concern that individuals who did not complete the four trials in a timely manner might have been discouraged by their performance on the first few trials. Accordingly, they may have exited the study, thereby producing a self-selection bias in that the participants who completed all four trials may have believed their dreams best matched the targets. A follow-up analysis (Watt & Valásek, 2015) indicated that the results of that study were indeed skewed toward cases in which participants’ dreams happened to match the targets, although after accounting for this self-selection bias the results trended toward significance.

Each dreamer submitted a report of the contents of any dreams that occurred over five nights in which they attempted to focus on the target video, which had at this point neither been viewed or even selected. After that report was submitted an independent judge ranked the similarity between that participant’s dreams and four video clips selected for the target pool using a random number generator based on atmospheric noise as the random source. The judge then completed a preplanned outcome measure, which was a ranking of the four clips in the target pool ranging from 1 (most similar to dream content) to 4 (least similar). After this ranking process was completed and submitted, the experimenters, who did not know the judge’s ranking results, used the random number generator to select one target video clip from the four ranked by the judge. The experimenters then sent a website link to this one video clip to the dreamer. The two earlier dream precognition studies were similarly well controlled, but those studies used only one dreamer who performed multiple trials (Krippner et al., 1972; Krippner, Ullman, & Honorton, 1971).

In sum, two well-controlled studies with statistically significant findings, two well-controlled studies with nonsignificant findings, and four controlled studies with potential for information confusion that yielded nonstatistically significant findings constitute too small a dataset from which to draw firm conclusions about whether dreams can reveal the content of upcoming unpredictable events. One mundane explanation for significant results is that they are drawn from a much larger pool of unpublished experiments with nonsignificant results (i.e., publication bias, a questionable research practice known to be pervasive across disciplines; e.g., Ioannidis, 2005; Pashler & Wagenmakers, 2012).

If the reader desires to further delve into the field, more detailed reviews of this topic are available elsewhere (Baptista, Derakhshani, & Tressoldi, 2015; Sherwood & Roe, 2013). Thus far, we may conclude that these data are insufficient for drawing conclusions. To better assess whether dreams can reveal veridical information about truly unpredictable future events, what is needed are repeated studies performed across multiple laboratories. Those studies should ideally use the same controlled group-study methods employed by Watt et al. (2015) with larger sample sizes and controls for self-selection bias, or with controlled single-participant methods (see Krippner et al., 1972; Krippner, Ullman, & Honorton, 1971).

A Note on Meta-Analysis

In many of the remaining sections, we report the results of meta-analyses. In the present dis-
cussion, we contend that meta-analysis is the best approach to determine whether a particular empirical effect replicates across studies, although it has been argued that meta-analysis is not an ideal tool for changing researchers’ minds (Ferguson, 2014). The Association for Psychological Science has included meta-analysis in its standards for building a cumulative discipline, as meta-analysis is the basis for evidence-based medicine (Cumming, 2014). The premise of meta-analysis is that a p value for any experiment may be above or below significance, even if the underlying effect is meaningful. This is because many uncontrolled factors can influence the result, especially in psychological experiments. When the average trend of the effect sizes is in the same direction, and the trend is statistically validated via a well-conducted meta-analysis, then we can be reasonably sure that the effect under consideration is meaningful (for updated reviews, see Lakens, Hilgard, & Staaks, 2016; Shamseer et al., 2015). In contrast, a single trial in a multi-trial experiment often provides a poor indicator of the overall average effect of interest. Similarly, a single successful or failed replication of any study should be regarded as tentative and with caution, providing a rationale for why (a) all attempted replications should be included in any meta-analysis and (b) we restrict our presentation of p values to those from meta-analyses and simply comment on the significance of only a few individual studies (assuming an alpha of 0.05 and two-tailed tests, with noted exceptions).

It important to note that although it is seldom discussed explicitly, the meta-analytic approach is based on an implicit assumption that the studies included report results that are not fraudulent, nor based on inappropriate multiple analyses (otherwise known as p-hacking, or fishing for a significant effect), and that all experimental results pertaining to the question under study were reported. Fraud, multiple analyses, and underreporting of negative results remain difficult to assess, but well conducted meta-analyses, such as those reviewed here, attempt to address these concerns.

Prospective meta-analyses are used in clinical medicine to further enhance the predictive value of meta-analysis. Caroline Watt of the University of Edinburgh is currently organizing such analyses for the parapsychological research community, based on her experience operating a study registry for precognition and other controversial experiments (Watt & Kennedy, 2015). In a prospective meta-analysis, the studies to be analyzed and the methods of analyses are preplanned. This approach offers a way to further avoid the biasing effects of potential fraud, multiple analyses, and underreporting of negative results. Prospective meta-analysis is an approach sympathetic to an argument made by Wagenmakers, Wetzels, Borsboom, van der Maas, and Kievit (2012), who suggest that all researchers should conduct confirmatory experiments with preregistered analyses and designs.

### Forced-Choice Conscious Precognition Tasks

A second approach to studying precognition is to test whether people can consciously access future information using a simple forced-choice task. In a forced-choice precognition task, a participant is offered two or more choices, one of which will randomly occur in the future, like the outcome of a coin toss. Participants perform multiple trials of the task, and the proportion of correct trials is the dependent variable. A meta-analysis of such experiments based on reports of 309 experiments published between 1935 and 1987 (Honorton & Ferrari, 1989) yielded a small overall effect size (Rosenthal ES = 0.02). Nevertheless, due to the high statistical power afforded by the many studies considered, it was statistically significant (Stouffer Z = 6.02, p < 1.1 × 10⁻⁹). Using Rosenthal’s failsafe estimate, the authors calculated that 14,268 unreported studies averaging a null effect would have been required to transform the database into one with an overall null effect. The size of that file-drawer estimate, in comparison with the number of laboratories studying precognition, suggested that selective reporting was an unlikely explanation for the observed effect. However, small effect sizes may also reflect consistent artifacts or methodological errors instead of a genuine effect.

Forced-choice, repeated-guessing precognition experiments began to fall out of favor in the 1990s for two reasons: (a) The tests were boring, encouraging participants to use conscious deliberation to try to outguess the next target, and (b) the tests did not resemble individuals’ spontaneous experiences of apparent precognition (Storm, Tressoldi, & DiRisio, 2010). So,
investigators began to gravitate toward two new types of precognition tasks: free-response conscious precognition tasks and implicit precognition tasks.

**Free-Response Conscious Precognition Tasks**

In a free-response conscious precognition task, the participant is asked to describe the contents of a randomly selected visual target that will be shown to the participant in the future. In a properly controlled precognition test, at the time the participant reports her or his impressions, no one (not even a computer) knows what the target will be. Only after the participant has recorded and submitted her or his impressions is a target randomly selected and presented to the participant.

Beginning in 1976, researchers from the Princeton Engineering Anomalies Research Laboratory performed multiple free-response precognition studies using this protocol. Dean of the School of Engineering and head of the Princeton Engineering Anomalies Research Lab, Robert Jahn, summarized those studies in a publication in the *Proceedings of the IEEE* (Jahn, 1982). He presented multiple “precognitive remote perception” experiments that resulted in highly significant results. A later analysis suggested that methodological weaknesses, including allowing investigators in one condition to freely choose targets according to their preferences, might have provided percipients with clues about the target in both precognitive trials and those in which the targets were chosen at the same time or before the percipients’ descriptions were recorded (Hansen, Utts, & Markwick, 1992). Thus, the overall findings were considered suspect until responses from the original experimenters made it clear that even after dropping the volitionally selected target trials, the data were still statistically significant (Dobyns, Dunne, Jahn, & Nelson, 1992; Nelson, Dunne, Dobyns, & Jahn, 1996). Further, taking into account only precognitive trials, data from a series using randomly selective targets located in the Chicago, Illinois, area as well as data from a series in which the percipient was located alone on a boat in the ocean could not be considered suspect because the percipient would have had no clues about the target (see “Chicago” and “Ocean” series; Jahn, 1982). Each of these precognitive remote viewing series were accurate significantly beyond chance, and neither suffered from the problems discussed by Hansen et al. (1992).

In support of the idea of using free-response instead of forced-choice tasks to examine precognition, Dunne and Jahn (2003) pointed out that trials using a truly free-response method produced larger effect sizes than those employing a partial-free-response (multiple-choice rating) approach (called feature importance discrimination option analysis). However, they performed this analysis as a post hoc attempt to explain what appeared to be a decline effect over time, and it is worth noting that some of the later trials that did not use feature importance discrimination option analysis still had generally smaller effect sizes than the original free-response trials.

May (2014) performed an experiment that exemplified a free-response precognition experiment with methods following the recommendations of Hansen et al. (1992). Three participants who had produced above-chance results in previous free-response tasks administered in the context of a classified government program, participated in a nonclassified experiment in which they were tasked with performing a total of 50 preplanned free-response conscious precognition trials. First, researchers asked a participant to describe a future target; this description was recorded. Then the experimenter made a “fuzzy-logic” judgment about that description prior to the random selection of a target image. To allow this judgment to be made without knowing the eventual target, researchers used a coding method that matched the elements in a participant’s description against a prearranged set of elements describing aspects of the target stimulus set (e.g., each potential target image had already been coded to indicate if the image depicted water, buildings, trees, etc.). After the judgment was completed, the target image and two additional decoy targets were randomly selected from a stimulus pool consisting of 300 images. Then, based on the number of matching elements between the previously coded de-
scriptions of each of the three target images, researchers calculated a “figure of merit” for the target image. After these calculations, the software randomly selected a target image, and the trial was judged to be a “hit” if the figure of merit for the target image was the highest figure of merit out of the three images. Finally, the target image was shown to the participant as feedback. This experiment produced 32 hits out of the 50 planned trials, a statistically significant hit rate based on a binomial test with a success rate of 33% expected by chance.

A significant result based on three participants is not worth mentioning if the experimenter’s goal is to determine whether an effect generalizes to the population at large. However, we understand that May’s (2014) experiment was simply a proof-of-principle demonstration that certain individuals can successfully predict future events that are specifically designed to be unpredictable by any ordinary means. Assuming that May’s results were not due to chance and that his report was accurate, his results disproved the null hypothesis. Independent attempts to replicate May’s experiment using the same participants would be both constructive and informative.

**Implicit Precognition Tasks**

The methods used in implicit precognition experiments may at first appear to be quite different from the methods already discussed, but they are similar in that implicit precognition tasks can be used to examine whether present actions are related to unpredictable future events. For example, we know that practicing or studying material that one desires to remember generally enhances recall of the studied material in the future. Daryl Bem (2011) conducted a clever experiment in which he reversed this practice effect in time to examine whether future practice might influence present performance on a recall task. Bem reported that the results of the first two of such tests revealed that participants were better at recall for words that they were going to practice in the future (Bem, 2011). This experiment and attempts to replicate its findings will be discussed in more detail below.

In recent decades, implicit precognition designs have become increasingly popular because they involve simple time-reversals of conventional tasks commonly used in experimental and social psychology. As such, these experiments are relatively easy to implement and they avoid the necessity of requiring participants to consciously guess future events, which researchers have long suspected impedes accessing future information (Carpenter, 2004, 2005).

Bem’s (2011) report on a series of nine implicit precognition experiments, all showing statistically significant or near-significant results, is perhaps the most widely discussed recent paper on the topic of implicit precognition. Bem was criticized for using 1-tail statistical tests, lack of clarity in reporting his methods, and potentially using multiple analyses to get the desired results (e.g., Alcock, 2011). Critics also suggested that a Bayesian approach should be used for all data analyses in psychology, instead of null-hypothesis significance-testing, with Bem’s report acting as the primary motivation for this suggestion (Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011). A response to that critique showed that after applying a Bayesian analysis with an appropriate prior distribution, the results were highly significant (Bem, Utts, & Johnson, 2011). Other authors also suggested that the analysis by Wagenmakers and others was flawed (Rouder & Morey, 2011). They reanalyzed Bem’s combined data using a Bayesian approach and found the results to be personally unconvincing, though it is clear that in a less controversial field the Bayes factor that they had calculated for Bem’s experiments with nonerotic stimuli would have been considered “moderate” to “strong” evidence (Goodman, 1999). Rouder and Morey (2011) stated at the end of the abstract of their report,

> There is some evidence...for the hypothesis that people can feel the future with emotionally valenced nonerotic stimuli, with a Bayes factor of about 40.

Although this value is certainly noteworthy, we believe it is orders of magnitude lower than what is required to overcome appropriate skepticism of ESP. (p. 682)

Others have been critical of the approach of combining different studies in this way, and they do not accept that there is strong enough evidence in favor of precognition based on these data (Wagenmakers, Wetzels, Borsboom, Kievit, & van der Maas, 2015).

In general, consensus exists among sceptics that many of the concerns raised about Bem’s research could have been avoided if Bem had
preregistered his experiments as confirmatory studies (see methodological recommendations, below). When there is too much freedom in the analytical steps that can be taken to confirm a hypothesized outcome, the probability increases that researchers will obtain their desired outcome. Preregistration of a confirmatory experiment, which includes specifying in advance the design and analytical strategies that will be adopted, constrains the experimenter to use only preplanned methodology, thereby mitigating or precluding confirmation bias in this or in any empirical field (Brandt et al., 2014).

Taken at face value, one of the most impressive findings from the Bem (2011) series of experiments was that practice on a list of words after a word-memory task was correlated with significant improvements in recall for the words that would subsequently be practiced. This was an implicit precognition test in that participants were not explicitly asked which words would be studied after they performed their word-memory test, but the test did use consciously deliberated responses (words typed by the participants) as the dependent variable. In contrast to these slower, more deliberate responses, other experiments in Bem’s report included designs that required relatively fast response times as the dependent variable.

The distinction between slow and fast responses is pertinent. The most statistically impressive results in Bem’s original report were from the two experiments investigating the time-reversed word-practice effect (Bem, 2011; Wagenmakers et al., 2011), but similar implicit precognition experiments, and some exact replications, using deliberate, slower responses (using so-called slow-thinking cognitive systems; e.g., Kahneman, 2011; Evans & Stanovich, 2013) have not produced reliably repeatable effects. It appears that experiments requiring fast responses (using so-called fast-thinking cognitive systems) have produced remarkably reliable and repeatable effects. This distinction was noted in a post hoc evaluation within a meta-analysis of 90 implicit precognition experiments conducted between 2000 and 2013, which included all known published and unpublished attempts to replicate Bem’s experiments (Bem, Tressoldi, Rabeyron, & Duggan, 2016). The authors used dual-process theory, as described by Kahneman (2011) as well as Evans and Stanovich (2013), to divide the experiments into fast versus slow-thinking—although, again, it is worth noting that this division was made by the authors after the authors were aware of each experiments’ results. According to that division, of the studies reported 61 were categorized as fast-thinking, including five categories: precognitive detection of reinforcement, precognitive avoidance of negative stimuli, retroactive priming, retroactive habituation, retroactive practice. The remaining 29 studies were categorized as slow-thinking, including the categories of retroactive facilitation of practice on recall and retroactive facilitation of practice on text reading speed.

Taken together, based on the original (not post hoc) analysis, the effect size for all 90 experiments combined was small but highly significant (Hedges’ $g = 0.09$; $p < 1.2 \times 10^{-10}$). The possibility of significant decline effects, that is, a reduction in the effect size of a given phenomenon over time, has been pointed out as a problem in parapsychology research (Hyman, 2010). If an effect declines to null over time, it is possible that the original effect was spurious, and as the quality of the experimental design and analyses improved, the effect would be understood as a fluke or an artifact. Analysis of the effect sizes for all 90 Bem-style experiments over time showed no significant decline effect (year vs. effect size: Pearson’s product-moment correlation: $-0.149$, $p > .161$). There was also little evidence that those results were due to multiple unreported analyses ($p$-hacking) or selective reporting (Bem et al., 2016). One concern that has been raised is that this meta-analysis included studies published prior to Bem’s (2011) study, but as pointed out in the meta-analysis, Bem started providing software for replications in 2000, the year of the first replication report included in this meta-analysis, so that concern seems unwarranted. In terms of fast- versus slow-thinking paradigms, the authors observed that all five categories of fast-thinking tasks independently produced significant results within each category, and the overall fast-thinking effect was highly significant ($z = 7.11$, $p < 6 \times 10^{-13}$). By contrast, slow-thinking tasks produced nonsignificant results in both experimental categories as well as overall ($z = 1.38$, $p > .15$).

Bem et al. (2016) suggested that the disparity between fast- and slow-thinking outcomes is
due to the former not suffering as much as the latter from conscious judgment or from alteration of the information obtained from nonconscious sources. An alternative explanation could be that many of the experiments requiring fast responses also used emotionally charged stimuli, which are more engaging than the stimuli used in the experiments requiring slower responses (Barušs & Rabier, 2014). If attention to precognitive information depends on motivation and engagement, which seems reasonable, then this explanation would both make intuitive sense and be empirically testable. A third possibility may be that implicit precognition follows a decay time-course similar to that obtained in studies of forced-choice conscious precognition (Honorton & Ferrari, 1989), such that tasks providing immediate feedback produce more accurate precognition than tasks providing delayed feedback.

A useful next step in a research program focusing on implicit precognition would be to continue to examine moderators of the effect. The analysis of moderators such as age, gender, and Big Five personality type (McCrae & John, 1992; Tupes & Cristal, 1961) could provide useful information about potential psychological, physiological and personality variables associated with precognitive effects, which might further shed light on whether precognition effects are due to subtle artifacts. In terms of potential mechanisms underlying purported precognitive effects, most fast-thinking paradigms currently conflate fast responses with immediate feedback. By teasing apart these two factors, researchers could determine if it is the speed of the response, the speed of the feedback, or both, contribute most to the precognitive effect. If the speed of the response matters most, this would suggest that deliberation over which response to select is what must be eliminated to access accurate precognition about future events. If instead the speed of the feedback matters most, this might suggest that precognition decays with a time course similar to decay of ordinary memory for past events. In general, variations of the implicit precognition paradigm offer many opportunities to study variables that interact with and modulate perception of future information, which in turn may offer insights into underlying physical mechanisms.

**Presentiment: A Physiological Measure of Implicit Precognition**

Outstanding predictive abilities, for example among elite athletes, may be based on highly efficient unconscious processing of sensory cues as well as exceptionally accurate predictive algorithms (e.g., Aglioti, Cesari, Romani, & Urgesi, 2008). This idea probably explains at least some aspects of superior performance. However, several decades of physiological evidence demonstrate that in humans and some animals, unconscious or semiconscious processes also seem to correctly anticipate unpredictable future events. This phenomenon has been called predictive anticipatory activity (Mossbridge, Tressoldi, & Utts, 2012; Radin, 2011), but a simpler moniker is presentiment (Radin, 1997; Radin & Pierce, 2015). As we describe below, presentiment may underlie superior performance in any sort of task that requires anticipation of noninferable future events.

In the first experiment that explicitly tested the presentiment hypothesis (Levin & Kennedy, 1975), contingent negative variation (CNV), a negative slow cortical potential that correlates with sensorimotor expectancy (Walter, Cooper, Aldridge, McCallum, & Winter, 1964), was used as a nonconscious physiological measure of precognition. Participants were asked to press a button as quickly as possible, but only after they saw a green light. After a red light, they were asked to withhold their button press. Electroencephalographic data showed a significant difference in CNV before the green light (when a motor action would be required), but not before a red light. Subsequent attempts to use CNV to measure presentiment failed to successfully replicate the effect (Hartwell, 1978, 1979).

Radin (1997) investigated the idea that emotional stimuli might be more likely to produce presentiment effects. He used heart rate, blood volume, and electrodermal activity as physiological measures. These measures were simultaneously recorded while participants viewed a randomized series of images. The protocol for these experiments included features that eventually became standard for most presentiment replication attempts, including the following: (a) Each trial consisted of an event that was either calm or emotional (e.g., viewing a picture...
of a tree vs. a plane crash), (b) the order of these events was completely randomized, (c) the experimental procedure provided no known sensory or statistical cues about the content or nature of the next event, and (d) the values of the physiological measures in a predefined period preceding each trial type were the dependent variables. This protocol revealed significantly different physiological changes preceding upcoming emotional events, as compared with calm events, with the difference emerging between 2 and 4 s, on average, prior to an event (Radin, 1997).

This basic result has been successfully replicated more than 20 times after Radin’s original study (e.g., McCraty, Atkinson, & Bradley, 2004; Radin, 2004; Radin & Lobach, 2007; Radin & Borges, 2009; Spottiswoode & May, 2003; Tressoldi, Martinelli, Semenzato, & Cappato, 2011). A statistically conservative meta-analysis of all known presentiment experiments following the above-described protocol concluded that presentiment is a repeatable effect (N = 26 studies; random effects: overall $ES = 0.21$, $z = 5.3$, $p < 5.7 \times 10^{-8}$, fixed effects: overall $ES = 0.21$, $z = 6.9$, $p < 2.7 \times 10^{-12}$), and that it occurs over a range of durations that appear to depend on the physiological system used to measure the effect (Mossbridge et al., 2012). An analysis of the meta-analytic data in Mossbridge et al. (2012) indicates that there is no significant decline effect in the findings over time (year vs. effect size: Pearson’s product-moment correlation $= -0.14$, $p > .49$). In this meta-analysis and also in a later review, potential problems were assessed including multiple unreported analyses ($p$-hacking), publication bias, and fraud. All three possibilities were rejected as viable explanations (Mossbridge et al., 2012; Mossbridge et al., 2014), at least partly because this meta-analysis tested a slightly different hypothesis than the original experimenters did. That is, almost all of the original experimenters tested a nondirectional hypothesis that physiological variables will differ consistently (in any direction) prior to emotional versus calm stimuli, whereas the authors of the meta-analysis tested a directional hypothesis that physiological variables will change in a consistent direction prior to emotional versus calm stimuli. The fact remains, of course, that $p$-hacking, publication bias, or fraud might not be detected in the data examined in any meta-analysis.

Of the various conventional explanations proposed to explain presentiment effects, the most plausible is the gambler’s fallacy. In this context the gambler’s fallacy would manifest as participants gradually becoming more anxious after a series of calm events because they would increasingly expect that an emotional event ought to occur soon. The fact that the underlying sequence of events is randomly determined, and independent of one another, is overpowered by inaccurate expectations about sequential randomness (e.g., Laplace, 1796/1951; Tversky & Kahneman, 1971). In fact, simulations have shown that a small presentiment-type outcome could potentially be explained by increases in physiological arousal due to a series of calm events that, by chance, just happened to precede an emotional event (Dalkvist, Westerlund, & Bierman, 2002; Wackermann, 2002). However, of the 26 experiments examined in the 2012 presentiment meta-analysis, 19 of the studies were performed by researchers who were aware of this potential confound and performed analyses to determine if the gambler’s fallacy bias could have reasonably explained the observed results. None of the researchers found convincing evidence for such an anticipatory bias (Mossbridge et al., 2012). In addition, a simulation of presentiment using a common expectation-bias test, a linear regression on the number of calm trials preceding an emotional one versus the physiological response preceding the emotional trial, found that 92% of the modeled presentiment effect remained unexplained even after removing experiments for which expectation bias detected with this test could potentially explain the results (Mossbridge et al., 2015).

Based on the results of the meta-analysis, the authors concluded that, barring widespread collusion among independent investigators, it appears that nonconscious access of future, unpredictable information is possible (Mossbridge et al., 2012). Of course, like the other experimental categories, research on presentiment effects has evoked debate (e.g., Mossbridge et al., 2015; Schwarzkopf, 2014). Schwarzkopf (2014) raised six points of concern, which were responded to in an exchange with Mossbridge and colleagues (2015). First, any meta-analysis is only as good as the data it considers. Meta-analyses, like any analysis, can be biased both in terms of which
articles are included and how the statistical analysis is performed (Franklin, Baumgart, & Schooler, 2014). Although Mossbridge and others performed a quality check and adopted a statistically conservative approach against their hypothesis whenever there was a choice of approaches, their meta-analysis might have been biased in some unspecified way. Schwarzkopf contended that the meta-analysis should have been broadened to include data from experiments not designed to test presentiment. In fact, Mossbridge et al. (2012) did analyze relevant psychophysiological data, but they did not include those results in their meta-analysis because the protocols did not use randomization of stimuli with replacement, an important feature of presentiment experiments intended to help reduce expectation bias. Schwarzkopf also expressed concern that because many of the presentiment experiments included in the meta-analysis used an uneven ratio of calming to arousing stimuli (often 2:1 or 3:1), which might have generated a response bias if participants figured out the stimulus ratio. Mossbridge and others pointed out that if this were the case, it would bias participants toward expecting calming stimuli, making presentiment more difficult to detect, not easier. Schwarzkopf then noted that clamping of the physiological trace to zero at a time prior to the stimulus could have produced the effects, but it was not clear how that could be the case given the analytical methods employed. Schwarzkopf further suggested that testing for expectation bias is generally performed by assuming that such biases would manifest linearly, but expectation biases might be nonlinear. Mossbridge et al. (2015) agreed. And finally, Schwarzkopf (2014) questioned if presentiment effects were biologically plausible. Mossbridge and others (2015) responded that if a presentiment effect was observed in these experiments, then there must be a natural explanation, even if one were not yet identified. Assessments of plausibility depend on current theoretical assumptions about what is or is not believed to be possible. Declaring a repeatable empirical effect to be implausible should arouse both caution and celebration; caution because expert intuitions about what is possible can be useful in identifying subtle confounds, artifacts, or experimenter biases, but also celebration because implausible effects can, at times, reveal entirely new, previously unimagined realms of knowledge.

Aside from these results in humans, a recent presentiment experiment in planaria (Alvarez, 2016) followed up a previous report of presentiment in birds by the same researcher (Alvarez, 2010). Experiments using lower level animals are useful in that they rule out common human forms of bias, such as the gambler’s fallacy. In the experiment with planaria, the worms were observed prior to a random number generator’s decision about whether a loud sound would be played. The experimenter measuring the behavior of the planaria was blind to whether the loud stimulus would be played. The planaria showed significantly more head movement (indicating stress or exploration) prior to when the loud stimulus was played as compared to times when it was not played. If researchers conduct a preregistered confirmatory experiment based on this finding, and if that experiment replicated the original effect, this would go a long way toward suggesting that biological mechanisms can use information about future events to influence present-time behavior.

Ideally, future presentiment experiments will continue to address what we believe to be the most outstanding concern: How do we know that the physiological effect we observe on any given trial is not caused by a delayed or latent response to the previous trial or trials? Most researchers performing psychophysiological experiments assume that trial randomization takes care of any such order effects and generally remain unconcerned about them, but those involved in presentiment research have been motivated to develop both analytical and experimental methods to determine whether expectation bias can explain presentiment (e.g., Dalkvist, Mossbridge, & Westerlund, 2014). As far as we know, the only guaranteed way to rule out order or expectation effects as an explanation for presentiment is to perform experiments in which each participant only experiences a single event, and in which comparisons between physiological measures preceding emotionally contrasting events are made across participants. The results of such a presentiment experiment, if they revealed a statistically significant presentiment effect, could strongly argue against any order- or expectation-based explanations for presentiment.
Recommendations for Future Investigations of Precognition

For decades, researchers interested in precognition and related phenomena routinely prespecified the numbers of participants and trials to eliminate “optional stopping.” This strategy can be especially effective when it is incorporated into a preregistered design in a prospective meta-analysis. More recently, researchers have recommended preregistration of experimental methods and analyses, especially for confirmatory experiments, to help ensure that researchers cannot change their planned experimental or analytical approaches after the study is underway (see http://www.koestler-parapsychology.psy.ed.ac.uk/TrialRegistry.html for several recently registered precognition experiments). Replications across multiple laboratories with data analyses performed by individuals who are unaware of either the hypothesis or the experimental manipulation(s) are particularly valuable. We also recommend that researchers attempting to perform a meta-analysis of precognition studies collaborate with a statistician who is unaware of the underlying hypothesis of the meta-analysis. This extra step would eliminate concerns about researcher biases affecting data analysis.

As for recommended next steps in precognition research, we contend that beginning to understand the psychological and physiological mechanisms underlying free-response precognition, implicit precognition, and physiological precognitive effects like presentiment should be the top research priority. Such a research program is especially important in light of the argument that potentially all other purportedly anomalous psychological phenomena, such as telepathy, clairvoyance, and psychokinesis, may be explainable via precognition (Marwaha & May, 2016). Another fruitful area is the examination of potential moderators of precognition such as age, personality type, mood, gender, and belief regarding the phenomenon. We may note that the use of prescreening to find participants skilled at the precognitive task of interest remains an important step for researchers attempting to gain insight into the mechanisms underlying precognition. Assuming that precognition is distributed unevenly in the population, like any other cognitive or physical talent, then any attempt to gain a deeper understanding of the mechanisms underlying these skills without studying participants who exhibit some skill is unlikely to be fruitful (Baruš & Mossbridge, 2016). Following these recommendations, future precognition research should work toward (a) estimating the robustness and repeatability of precognitive effects; (b) gaining a clearer understanding of the relationships, if any, among different types of precognitive tasks; and (c) shedding light on underlying mechanisms. More mechanistically inspired recommendations for future work are offered below (see Summary).

Potential Psychological Mechanisms of Precognition

Based on the experiments discussed above, it appears that nonconscious mental processes are largely responsible for precognition (Stanford, 2015). Historically, scientists and the lay public have resisted the idea that a part of the mind functions in a way that is completely beyond conscious awareness. Subliminal perception is now openly discussed in most psychology texts and courses, but for many decades the idea of nonconscious cognitive processes was regarded as ridiculous (Kihlstrom, Barnhardt, & Tataryn, 1992; Hassin, 2013). Indeed, the idea that nonconscious processing associated with dreaming has any utility has only recently gained scientific respectability after robust evidence showed that dreaming influences learning (e.g., Bob & Louchakova, 2015; Hobson & Pace-Schott, 2002; Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994; Llewellyn & Hobson, 2015; Stickgold, Hobson, Fosse, & Fosse, 2001). Further, only in the last 15 years have most researchers in psychology and neuroscience acknowledged that nonconscious or implicitly accessed cognitive processes can sometimes assist in recalling memories, making choices, and solving complex problems better than conscious cognitive processes (Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Dijksterhuis, Bos, Nordgren, & van Baaren, 2006; Voss, Baym, & Paller, 2008; Voss, Lucas, & Paller, 2012). In other words, evidence continues to accumulate that our intuitions about the nature, scope, and abilities of our conscious and unconscious awareness are not necessarily accurate.

The frailty of our intuitions becomes especially clear when we are confronted by data suggesting unconscious awareness of future
events. Not only is a persuasive theoretical explanation for precognition unavailable at this point, but we do not even know whether we should be attempting to identify one mechanism or many. Perhaps the time span between the precognition and the event is important; perhaps there is one mechanism for behavioral precognition (like precognitive remote viewing and precognitive dreaming, with time frames on the order of days) and a different mechanism for physiological precognition (like presentiment, with time frames on the order of seconds). Another unresolved question is whether precognition is achieved via accessing information about probable or actual future events, or by events in the future “influencing” or constraining events in the past.

Summary

The full epistemological and ontological consequences of time-reversed influences are not yet clear, but one implication is that the experimental sciences may soon be faced with a troubling dilemma: Time-reversed effects, if they exist, cannot be prevented by any currently known experimental controls. As we have seen in this review, several classes of experiments have demonstrated time-reversed anomalies under tightly controlled protocols. Accordingly, our most cherished epistemologies may be unavoidably influenced by future outcomes. We may take comfort in the likelihood that the magnitudes of these influences are probably small, but in some disciplines, especially domains like the life sciences and experimental psychology in which thousands of variables influence the observed effects, time-reversed effects may fundamentally affect the interpretation of results.

Such speculative implications, of course, can be considered scientific heresies of the first order. But if positive empirical evidence continues to accumulate, especially if the methodological recommendations suggested by ourselves and others are followed, then a time may come when we are forced to think the unthinkable. Indeed, the implications of retrocausation are so remote from engrained ways of thinking that the first reaction to this line of research is that it must be flawed. The second reaction may be horror that it represents a previously unaccepted fact about reality.

To better understand the nature of precognition, we need to study the relationships between nonconscious processes, conscious processes, and how time unfolds in the physical world. Several avenues of inquiry that may lead to greater understanding include examining the circumstances under which nonconscious processes share information with conscious processes about imminent events, determining how a conscious decision to receive information about a future event influences one’s ability to accurately perceive that information, collaborating with physicists to study how nonconscious and conscious processes might interact with events as they unfold in time, and examining how time is perceived during alterations in consciousness.

A recent “taxonomy of prospection” delineates four broad categories in which most skills related to prospection seem to fall: simulation, prediction, intention, and planning (Szpunar, Spreng, & Schacter, 2014). Based on the data reviewed here, it seems to us that precognition may eventually be considered just one of several forms of prediction that have evolved to enhance our survival. A handful of neuroscientists, psychologists, and physicists are examining precognition with this idea in mind, and some have published their results (e.g., Bem, 2011; Franklin et al., 2014; Mossbridge, Tresoldi, & Utts, 2012). However, due to a lack of awareness of this line of research in mainstream academia, such efforts are vastly underfunded. That is a pity because the ability to gain information about future events could potentially lead to major advances in both psychological and physical theories, as well as a host of pragmatic applications. We join others (Franklin et al., 2014) in supporting efforts to increase funding for precognition research.

References


PRECOGNITION AS A FORM OF PROSPECTION


Voss, J. L., Baym, C. L., & Paller, K. A. (2008). Accurate forced-choice recognition without aware-


Received November 9, 2015
Revision received February 14, 2017
Accepted February 26, 2017