

Feeling the past: The absence of experimental evidence for anomalous retroactive influences on text processing

Matthew J. Traxler · Donald J. Foss · Ruchira Podali · Megan Zirnstein

Published online: 20 July 2012
© Psychonomic Society, Inc. 2012

Abstract In two self-paced reading experiments, we investigated the hypothesis that information moves backward in time to influence prior behaviors (Bem *Journal of Personality and Social Psychology* 100:407–425, 2011a). In two of Bem's experiments, words were presented after target pictures in a pleasantness judgment task. In a condition in which the words were consistent with the emotional valence of the picture, reaction times to the pictures were significantly shorter, as compared with a condition in which the words were inconsistent with the emotional valence of the picture. Bem *Journal of Personality and Social Psychology* 100:407–425, (2011a) interpreted these results as showing a “retroactive priming” effect resulting from precognition. To test the precognition hypothesis, we adapted a standard repetition priming paradigm from psycholinguistics. In the experiments, participants read a set of texts. In one condition, the participants read the same text twice. In other conditions, participants read two different texts. The precognition hypothesis predicts that readers who encounter the same text twice will experience reductions in processing load during their first encounter with the text. Hence, these readers' average reading times should be shorter than those of readers who encounter the target text only once. Our results indicated that readers processed the target text faster the second time they read it. Also, their reading times decreased as their experience with the self-paced reading

procedure increased. However, participants read the target text equally quickly during their initial encounter with the text, whether or not the text was subsequently repeated. Thus, the experiments demonstrated normal repetition priming and practice effects but offered no evidence for retroactive influences on text processing.

Keywords Context effects · Discourse processing · Language comprehension · Psycholinguistics · PSI · Parapsychology · ESP · Precognition · Retrocausation

Ideas should not be judged scientific or unscientific, true or false, on the basis of their origins. Truths may come from sources that are quite unreliable, and false theories may come from the most trustworthy persons applying the most rigorous methods. (Paul D. Allison¹)

Belief in paranormal phenomena is widespread in the general population. These beliefs are reinforced by frequent uncritical reporting of parapsychological claims in the popular media (Stanovich, 1998). It should come as no surprise, then, that some professionally trained academic psychologists also express belief in psychic phenomena. In fact, psychology as a discipline once had a very close relationship with paranormal research (Allison, 1979; Hansel, 1989). Parapsychology research has been conducted at one time or another, often but not always in psychology departments, at Harvard, Cambridge, Princeton, Stanford, Columbia, Duke, Cornell, Syracuse, the University of Arizona, the University of Virginia, and the University of Edinburgh, among others. Many prominent psychologists, including William James and one of his successors, William McDougall (who held the William James Chair of Psychology at

M. J. Traxler (✉) · M. Zirnstein
Department of Psychology, University of California, Davis,
1 Shields Avenue,
Davis, CA 95616, USA
e-mail: mjtraxler@ucdavis.edu

D. J. Foss · R. Podali
Department of Psychology, University of Houston,
Houston, TX, USA

¹ Allison, 1979.

Harvard University), engaged in parapsychological research.² Although psychology as a discipline has largely abandoned research on the paranormal as a topic of serious inquiry,³ mainstream psychology journals continue to publish research articles, commentaries, and meta-analyses on aspects of parapsychology every once in a while. Many of these articles purport to provide support for a variety of paranormal abilities, which are sometimes collectively termed *psi*. Paranormal abilities are said to include the ability to influence objects or physical processes without physical interaction, through mental effort alone (*psychokinesis*), the ability to detect events at a distance, without the possibility of perceptual input or sensory processing (*remote viewing* or *clairvoyance*), and the ability to predict future events via processes that go beyond normal processes of inference and deduction (*precognition* or *premonition*).

Our chief purpose in conducting the research reported here was to test recently advanced claims regarding the existence of precognition in representative samples of college-age participants (Bem, 2011a). In prior experiments (reviewed below), standard social and cognitive experimental paradigms have been adapted in such a way that their outcomes have the potential to show the influence of future events on participants' current responses. We have attempted a conceptual replication. More specifically, we generated from the precognition hypothesis predictions for outcomes in a standard psycholinguistic priming paradigm (e.g., Long, Oppy, & Seely, 1997; see Traxler, 2012; Traxler & Gernsbacher, 2006). There is a strong temptation to simply ignore experimental work purporting to show evidence of paranormal phenomena. However, given the prominence of recent work of this type, and given the seriousness with which it was conducted and presented, it will be useful to explore whether retroactive influence can be observed in a novel cognitive domain, that of text processing.

In the present experiments, we attempted to replicate prior retroactive priming effects in a similar participant sample, using a closely related but not identical technique. Such converging evidence, if obtained, would greatly strengthen the case for the existence of paranormal abilities in college-age participants. Essentially, we are asking whether retroactive influences apply to language processing, since they have been claimed to apply to memory, and perception of emotionally salient visual stimuli.

In an article entitled, "Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect," Daryl Bem (2011a) purported to show evidence for precognition in college students who were representative of the general student body at Cornell

University. In a series of nine experiments, standard social psychology and cognitive psychology experimental paradigms were adapted to test the hypothesis that information from the future could flow backward in time to affect participants' current responses: "The experimental procedures were based on simple, well-established psychological effects that would be familiar to most readers" (p. 420). Some of these experiments involved a study–test paradigm, some involved making pleasantness judgments of pictures, and some involved attempting to predict the future location of a picture.

In the priming experiments, participants judged the emotional valence of pictures (Bem, 2011a, Experiments 3 and 4). In a normal affective priming experiment, participants read a prime word that has either a positive or a negative emotional tone (e.g., *pretty*, positive valence; *ugly*, negative valence; Bargh & Ferguson, 2000). They are then exposed to a picture that depicts a positive or negative scene and are asked to press a button indicating their judgment about the valence of the picture. In standard priming experiments of this type, judgments that are consistent with the valence of the prime (positive picture following positive word; negative picture following negative word) are completed faster than judgments that are incongruent with the prime word. In Bem's (2011a) Experiments 3 and 4, the same task was used, but the prime words appeared after each picture judgment response was completed. Using different outlier-trimming schemes, Experiment 3 showed between 15- and 24-ms reductions in response time when the prime word was congruent with the picture, as compared with when it was incongruent. Experiment 4 showed a 17- to 27-ms savings, again depending on which exact outlier removal procedure was applied to the data. These results were interpreted as showing that information moves backward in time, such that future experience can influence current performance.

Experiment 1a

The experiments reported here were designed as a conceptual replication of Bem's (2011a) experiments. As Bem noted (2011a), "the major empirical challenge . . . is to provide well-controlled demonstrations of psi that can be replicated by independent investigators" (p. 407). We implemented a standard repetition priming paradigm that will be familiar to readers who follow the psycholinguistics literature, even casually. In a standard repetition priming experiment, participants are exposed to a linguistic stimulus at time 1, and their response to another instance of that same stimulus is recorded at time 2. A variety of prime–target characteristics can be manipulated to assess a variety of cognitive processes, but our focus here is on repetition priming effects. Likewise, a variety of tasks and dependent measures can be used to assess participants' reaction to the

² Readers who are interested in the history of parapsychology research in psychology and related disciplines should consult Hansel's book *ESP and Parapsychology Revisited* for a thorough, engaging review.

³ In fact, many universities are hesitant to accept funding from outside sources for paranormal research (Wallis, 1985).

prime stimulus. Here, we used self-paced reading, a common research technique. If information flows backward in time, exposing participants to an identity prime at time 2 should speed their response to another instance of the same stimulus at time 1. If obtained, this pattern would extend the range of stimuli to which retroactive influence applies and would provide converging evidence for precognition. Failure to replicate retroactive influence effects would cast doubt upon precognition as an influence on performance in language-processing tasks.

Method

Participants Forty-eight undergraduates from the University of California, Davis participated in partial fulfillment of a course requirement. All participants were native English speakers with normal hearing and vision.

Stimuli and procedure The experimental stimuli consisted of two stories adapted from Wikipedia. One of the texts was on the subject of paranormal research, and the other was on the subject of memory research. The texts are available upon request from the corresponding author (mjtraxler@ucdavis.edu). The critical experimental text consisted of 1,009 words.

The participants' task was to read and understand the texts. Each participant read two texts. Participants were randomly assigned to one of three groups. In the *ESP-ESP* group, participants first read the text about ESP research. After finishing that text, they read the same text again. This condition produces two sets of results. The first result, the *ESP-target* result, is based on the first encounter with the text. The second result, the *repetition priming* result, is based on the second encounter with the text. The *ESP-target* result provides an estimate of effects of precognition on reading time. If information from the second reading of the text influences processing during the first encounter, participants should have read the ESP text faster than normal in the *ESP-target* condition. The repetition priming result provides an estimate of the savings obtained when participants reread the same text.

In the *ESP-baseline* group, participants first read the text about ESP. After finishing that text, they read a text about memory research. In this condition, no precognitive effects should be possible, because the ESP text is not repeated. Hence, the *ESP-baseline* condition gives us an estimate of "normal" reading time for the ESP story.

In the *practice control* group, the participants read the text about memory research first, followed by the text about ESP research. The practice control condition provides us an estimate of savings provided by practice with the self-paced reading technique itself.

Self-paced reading procedure Participants were tested individually in a quiet room. They were instructed to read at a normal, comfortable pace in a manner that would enable them to answer comprehension questions. We did not actually ask readers comprehension questions after they finished reading the two texts, however. The texts were presented with a self-paced moving window procedure running on a desktop PC computer. Each trial began with a series of dashes on the computer screen in place of the letters in the words. Any punctuation marks appeared in their exact position throughout the trial. The first press of the space bar replaced the first set of dashes with the first word in the text. With subsequent space bar presses, the next set of dashes was replaced by the next word, and the preceding word was replaced by dashes. The computer recorded the time from when a word was first displayed until the next press of the space bar.

Results and discussion

Figure 1 presents mean reading time per word by condition.

Tests for precognition and practice effects We conducted separate by-participant and by-item 1×3 (condition: *ESP-target* vs. *ESP-baseline* vs. *practice control*) ANOVAs. These ANOVAs revealed no differences in mean reading times between the three conditions in the by-participants analysis, $F_1(2, 45) < 1$, $MSE = 13,314$, n.s.. However, the by-items analysis did show an effect of condition, $F_2(2, 2016) = 10.8$, $MSE = 3,585$, $p < .001$. The *ESP-target* and *ESP-baseline* conditions did not differ, $t_2(1008) < 1$, n.s.]. The *ESP-target* and *practice control* conditions did differ in the by-items analysis, $t_2(1008) = 4.00$, $p < .001$, as did the *ESP-baseline* and *practice control* conditions, $t_2(1008) = 3.57$, $p < .001$.

The null hypothesis significance test provides an estimate of how likely it is that the observed p -value will be obtained, if the null hypothesis is true (i.e., that two samples are drawn

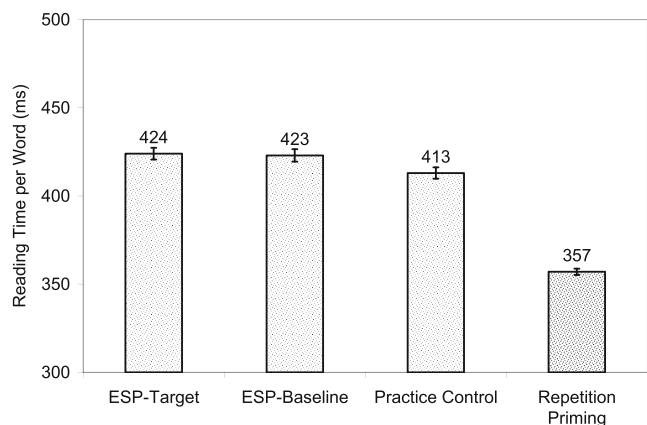


Fig. 1 Mean reading time by condition for Experiment 1a

from the same underlying distribution). However, obtaining a p -value that does not reach the standard level of significance does not, by itself, justify the conclusion that there is no actual difference between two conditions (i.e., the absence of evidence is not the same thing as the evidence of absence.) Thus, finding a nonsignificant p -value in a contrast of the ESP–target and the ESP–baseline conditions does not allow us to conclude that there was no benefit of future repetition in the ESP–target condition.⁴ In contrast to null hypothesis testing, a Bayes factor analysis allows one to quantify the likelihood of a pattern of results if the null hypothesis is true, as compared with the likelihood of those same results if there really were a difference between two conditions (Rouder, Speckman, Sun, Morey, & Iverson, 2009). When we conducted a Bayes factor analysis based on the contrast between the ESP–target condition and the ESP–baseline condition, it indicated that the obtained results were nearly 40 times more likely under the null hypothesis than under an hypothesis where the ESP–target and ESP–baseline conditions differ ($t = 0.417$, $N = 1,009$, posterior probability of null hypothesis/probability of alternative = 36.5).⁵ According to Jeffreys (1961), this outcome constitutes "very strong" evidence in favor of the null hypothesis over the alternative.

The results do indicate that having practice with the self-paced reading technique produces shorter reading times, even when the preceding text is unrelated to the text from which reading times were drawn.

Tests for repetition priming To test for repetition priming, we contrasted the ESP–target condition with the repetition priming condition. Reading times were significantly shorter in the repetition priming condition, relative to the ESP–target condition, $t_1(15) = 5.04$, $MSE = 17.8$, $p < .001$; $t_2(1008) = 22.8$, $MSE = 2.94$, $p < .0001$. Thus, the results show that reading the ESP text once leads to substantial reductions in reading time the second time the same text is encountered.

Before interpreting these results further, we turn to the results of Experiment 1b.

⁴ "(C)onventional significance tests do not allow the analyst to state evidence for a null hypothesis. If an invariance holds, even approximately, then the best-case significance test outcome is a failure to reject, which is interpreted as a state of ignorance." (Rouder et al., 2009, p. 226)

⁵ In this analysis, we set the prior odds of the null and alternative hypotheses conservatively at 1. In truth, the prior probability of the null hypothesis is likely to be much greater than the precognition hypothesis. Hence, our Bayes factor results almost certainly underestimate the posterior odds of the null hypothesis relative to the precognition hypothesis. If we had assumed that the prior probability of the null was twice as high as the precognition hypothesis, the Bayes factor would have been greater than 70.

Experiment 1b

Experiment 1b was identical in every respect to Experiment 1a, except that it involved a different sample of college students from the University of Houston. Experiment 1b provided us the opportunity to assess whether key results from Experiment 1a would replicate, as well as providing another opportunity for precognitive effects on reading time to emerge.

Method

Participants Participants were 60 undergraduate students from the University of Houston.

Stimuli and procedure The stimuli and procedure were identical to those in Experiment 1a.

Results and discussion

Figure 2 presents mean reading time per word by condition.

Tests for precognition and practice effects The 1×3 (condition: ESP–target vs. ESP–baseline vs. practice control) ANOVAs revealed robust differences in mean reading times between the three conditions in the by-participants and by-items analyses, $F_1(2, 57) = 6.03$, $MSE = 21,145$, $p < .005$; $F_2(2, 2016) = 1,599$, $MSE = 4,179$, $p < .0001$. The ESP–target and baseline conditions did not differ in the by-participants analysis, t_1 , n.s., but they did differ according to the by-items analysis, $t_2(1008) = 2.68$, $p < .01$. This result indicates that participants read the ESP text faster in the baseline condition than in the ESP–target condition, contrary to the precognition hypothesis, which predicts the opposite outcome. More likely, the result reflects minor random variation in reading speed among the participants

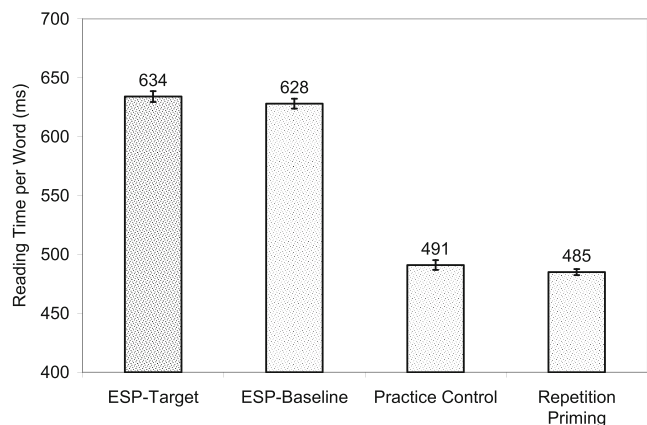


Fig. 2 Mean reading time by condition for Experiment 1b

assigned to the different conditions. This variation is not enough to influence the outcome of the by-participants analysis, which has considerably less power than the by-items analysis.

Despite the positive result in the by-items analysis (positive, that is, but in the wrong direction according to the precognition hypothesis), the Bayes factor analysis for this contrast shows that the posterior odds of obtaining this result is about the same under the null hypothesis as it is under an alternative where the baseline condition produces shorter reading times than the ESP–target condition (posterior odds of the null hypothesis/alternative = 1.12). Thus, the Bayes factor analysis supports the conclusion that the two conditions really do not differ.

Reading times in the ESP–target condition were reliably longer than those in the practice control condition, $t_1(39) = 3.09$, $p = .01$; $t_2(1008) = 42.5$, $p < .0001$. This result presumably reflects ordinary adaptation to the self-paced reading procedure.

Finally, reading times in the ESP–baseline condition were reliably longer than reading times in the practice control condition, $t_1(39) = 2.93$, $p < .05$; $t_2(1008) = 46.5$, $p < .0001$.

As in Experiment 1a, these outcomes do not provide any evidence that reading a text the second time has any influence on reading times during the first encounter with that text. They do indicate that having practice with the self-paced reading technique produces shorter reading times, even when the preceding text is unrelated to the text in question.

Tests for repetition priming To test for repetition priming, we contrasted the ESP–target condition with the repetition priming condition. Reading times for the second encounter with the ESP text were significantly shorter than those for the first encounter, $t_1(19) = 8.34$, $MSE = 17.5$, $p < .001$; $t_2(1008) = 40.3$, $MSE = 3.71$, $p < .0001$.

General discussion

In Experiments 1a and 1b, one group of participants read a text about ESP followed by the identical text. Another group read the same ESP text followed by a different text about memory research. The last group was assigned to the practice control group, which read the ESP text after they read the memory text. In both subexperiments, participants exhibited strong and significant repetition priming effects upon encountering the ESP text for the second time. In both subexperiments, participants in the ESP–target condition read the ESP text just as slowly as did participants in the ESP–baseline condition. In fact, in both cases, mean reading time per word was within 6 ms (1 ms/word in Experiment

1a; 6 ms/word in Experiment 1b). In Experiment 1a, the 1-ms difference was not significant in either the by-participants or the by-items analysis. In Experiment 1b, the 6-ms difference was significant in the by-items analysis, but not in the by-participants analysis. Whether there is a true difference in reading time between the two conditions, the numerical difference runs counter to the retroactive influence hypothesis (i.e., reading times were longer in the ESP–target condition than in the ESP–baseline condition).

Neither of the experiments produced evidence for effects of the second encounter with the ESP text on processing during the first encounter, anomalous or otherwise. One might conjecture that this null effect could be attributed to a lack of sensitivity or statistical power. There is substantial evidence against this possibility, however. A Bayesian t -test on the data from Experiment 1a showed that the observed outcomes were nearly 40 times more likely to occur under the null hypothesis than under an alternative where baseline and ESP–target reading times differed. The standard rubric for interpreting Bayes factors classifies this as "very strong" evidence in favor of the null hypothesis (Jeffreys, 1961). Experiment 1b offers even stronger evidence against the precognition hypothesis, since the marginal numerical result runs opposite to the result predicted by the precognition hypothesis (although the Bayesian t -test indicated that the results were nearly as likely to occur under the null hypothesis). Second, using standard procedures for computing and interpreting results in the psycholinguistic tradition, we detected mean differences between conditions as small as 10 ms. This difference is, in fact, smaller than the 15- to 27-ms effects found in previous retroactive priming experiments that had much larger numbers of participants than did ours (Bem, 2011a, Experiments 3 and 4). Thus, our experiments appeared to be just as powerful and just as sensitive to numerically small reaction time effects as previous experiments that have been taken as evidence for precognition. We could adopt a more liberal standard of evidence by accepting a statistically significant result in either the by-participants or the by-items analysis. Even if we did this, however, there is no instance where mean reading times for the ESP–target condition were shorter than reading times for the ESP–baseline condition.

These results contrast sharply with Bem's (2011a) study, in which 9 of 10 experiments produced statistically significant effects. What accounts for the discrepancy? The most likely possibility is that precognition did not influence the results of either our experiments or Bem's (2011a). Recent work in Bayesian statistics suggests that the statistical evidence from Bem's (2011a) study is not as strong as it first appears (Rouder & Morey, 2011). Furthermore, Ritchie and colleagues recently attempted to replicate the original retroactive facilitation of recall effects, using an extremely

rigorous preregistered experiment technique. None of Ritchie and colleagues' experiments replicated the effects reported in the previous study (Ritchie et al. 2012). An additional recent article reported that the original study showed evidence of *publication bias* (Francis, 2012). That is, more statistically significant outcomes were reported than would be expected given the reported effect sizes and statistical power of the individual experiments. The presence of publication bias may itself reflect reporting of exploratory analyses as confirmatory, otherwise known as *prospecting* (Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011; see also Diaconis, 1978). In a situation where precognition does not have any real effect on outcomes, reported positive results may indeed reflect the "long arm of chance." (For additional discussion of statistical methods relating to investigation of parapsychological hypotheses and related issues, see Bem, 2011b; Hyman, 1994, 2010; Kruschke, 2011; Lebel & Peters, 2011; Loftus, 1996; May, Utts, & Spottiswoode, 1995; Storm & Ertel, 2001; Storm, Tressoldi, & Di Risio, 2010; Utts, 1991; Wetzels et al., 2011).

One might object to this line of reasoning by proposing that the present experiments operated on a different time scale than prior affective priming studies (Bem, 2011a). In our case, several minutes might elapse between the first encounter with the ESP text and the second. One might therefore attribute the difference in results between the two studies to the difference in elapsed time between the two critical stimuli. While this is certainly a possibility, there are theoretical and empirical considerations that argue against it.

First, in terms of theoretical issues, a generic account of paranormal abilities posits that information flow is liberated from the constraints of time and space. "Paranormal processes such as telepathy and precognition reveal instantaneous events transcending boundaries of space and time" (Bourne, 2008, p. 79; see also Jones, 2009). In some of Rhine's early work involving "senders" and "receivers," a wide variety of temporal conditions were tested, and no reliable differences were obtained between conditions (Hansel, 1989). Furthermore, no theory of paranormal abilities has been expounded that would rule out the possibility of retroactive influence lasting for several minutes. Thus, there is no theory of precognition that would predict a null result in our experiments based on the amount of time that elapsed between the two critical stimuli. In addition, one might think that structured information, such as a discourse representation, would be more resistant to retroactive decay or interference. Our stimuli were arguably richer in information content than were Bem's (2011a) stimuli. As such, their retroactive influence could easily transcend greater amounts of time. Minimally, no extant theory or findings contradict this claim. Finally, previous retroactive facilitation of recall experiments involved time scales of several minutes, similar to the present experiments. If doubt

remains, list priming methods involving lexical decision or naming could easily be adapted to investigate the same hypotheses.

While the results of Experiments 1a and 1b are fully consistent on the most critical comparisons, the results were not completely identical. The one substantive difference in the results from Experiments 1a and 1b was in the magnitude of the self-paced reading practice effects, which were much larger in Experiment 1b than in Experiment 1a. This outcome may reflect differences in susceptibility to fatigue or overall degree of tolerance for reading esoteric material between the two groups of participants.

Conclusions

Our experiments showed that processing a text at time 1 had strong influences on processing of that same text at time 2. The experiments also showed that general experience with a reading task at time 1 speeded the execution of that task at time 2. These results are compatible with those of many other studies and with all prevailing psycholinguistic accounts of language processing and reading. By contrast, the experiments provided no evidence for retroactive influences on reading performance. The results are therefore consistent with a simple summary of the standard view of mind: Information that does not exist in the mind at time T cannot influence performance at time T , even if that information will be created in that mind at a later time.

Author Notes The authors thank Dr. Reed Hunt, Dr. Henry L. Roediger III, and an anonymous reviewer for helpful comments on a previous draft of the manuscript.

References

- Allison, P. D. (1979). Experimental parapsychology as a rejected science. In R. Wallis (Ed.), *On the margins of science: The social construction of rejected knowledge* (pp. 271–291). Keele: University of Keele.
- Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.
- Bem, D. J. (2011a). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407–425.
- Bem, D. J. (2011b). Must psychologists change the way they analyze their data? *Journal of Personality and Social Psychology*, *101*, 716–719.
- Bourne, E. J. (2008). *Global shift: How a new worldview is transforming humanity*. Oakland: New Harbinger.
- Diaconis, P. (1978). Statistical problems in ESP research. *Science*, *201*, 131–136.
- Francis, G. (2012). Too good to be true: Publication bias in two prominent studies from experimental psychology. *Psychonomic Bulletin and Review*, *19*, 151–156.

- Hansel, C. E. M. (1989). *The search for psychic power: ESP and parapsychology revisited*. Amherst: Prometheus.
- Hyman, R. (1994). Anomaly or artifact? Comments on Bem and Honorton. *Psychological Bulletin*, *115*, 19–24.
- Hyman, R. (2010). Meta-analysis that conceals more than it reveals: Comment on Storm et al. *Psychological Bulletin*, *136*, 486–490.
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press.
- Jones, M. D. (2009). *Modern science and the paranormal*. New York: Rosen Publishing Group.
- Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, *6*, 299–312.
- LeBel, E., & Peters, K. R. (2011). Fearing the future of empirical psychology: Bem's (2011) evidence of Psi as a case study of deficiencies in modal research practice. *Review of General Psychology*, *15*, 371–379.
- Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science*, *5*, 161–171.
- Long, D. L., Oppy, B. J., & Seely, M. R. (1997). Individual differences in readers' sentence and text-level representations. *Journal of Memory and Language*, *36*, 129–145.
- May, E. C., Utts, J. M., & Spottiswoode, S. J. P. (1995). Decision augmentation theory: Toward a model of AMP. *Journal of Parapsychology*, *59*, 195–220.
- Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem's 'Retroactive Facilitation of Recall' effect. *PLoS One*, *7*, 1–5.
- Rouder, J. N., & Morey, R. D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin and Review*, *18*, 682–689.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review*, *16*, 225–237.
- Stanovich, K. E. (1998). *How to think straight about psychology*. New York: Longman.
- Storm, L., & Ertel, S. (2001). Does psi exist? Comments on Milton and Wiseman's (1999) meta-analysis of ganzfeld research. *Psychological Bulletin*, *127*, 424–433.
- Storm, L., Tressoldi, P. E., & Di Risio, L. (2010). Meta-analysis of free-response studies, 1992–2008: Assessing the noise-reduction model in parapsychology. *Psychological Bulletin*, *136*, 471–485.
- Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*. Boston: Wiley-Blackwell.
- Traxler, M. J., & Gernsbacher, M. A. (2006). *The handbook of psycholinguistics* (2nd ed.). Amsterdam: Elsevier.
- Utts, J. (1991). Replication and meta-analysis in parapsychology. *Statistical Science*, *6*, 363–403.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*, 426–432.
- Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855t-tests. *Perspectives on Psychological Science*, *6*, 291–298.