This article presents a comprehensive integration of current experimental evidence and theories about so-called parapsychological (psi) phenomena. Throughout history, people have reported events that seem to violate the common sense view of space and time. Some psychologists have been at the forefront of investigating these phenomena with sophisticated research protocols and theory, while others have devoted much of their careers to criticizing the field. Both stances can be explained by psychologists’ expertise on relevant processes such as perception, memory, belief, and conscious and nonconscious processes. This article clarifies the domain of psi, summarizes recent theories from physics and psychology that present psi phenomena as at least plausible, and then provides an overview of recent/updated meta-analyses. The evidence provides cumulative support for the reality of psi, which cannot be readily explained away by the quality of the studies, fraud, selective reporting, experimental or analytical incompetence, or other frequent criticisms. The evidence for psi is comparable to that for established phenomena in psychology and other disciplines, although there is no consensual understanding of them. The article concludes with recommendations for further progress in the field including the use of project and data repositories, conducting multidisciplinary studies with enough power, developing further nonconscious measures of psi and falsifiable theories, analyzing the characteristics of successful sessions and participants, improving the ecological validity of studies, testing how to increase effect sizes, recruiting more researchers at least open to the possibility of psi, and situating psi phenomena within larger domains such as the study of consciousness.

Keywords: parapsychology, psychical research, psi, meta-analysis, anomalous cognition

People in all walks of life have reported events that seem to violate the current common sense view of space and time, from dreams that seem to ostensibly predict a noninferable, dramatic event, to the more mundane assertion by a former prime minister of Sweden that he can sense when his wife is about to call him (Thunberg, 2006). In various surveys, majorities of respondents have endorsed a belief in such phenomena, which may have a noticeable impact on their lives (Watt & Tierney, 2014). In the last few years, parapsychology (psi) research has appeared in major psychology journals (e.g., Bem, 2011; Storm, Tressoldi, & Di Risio, 2010a, 2010b), and comprehensive reviews of the evidence for and against psi have been published (Cardeña, Palmer, & Marcusson-Clavertz, 2015; May & Marwaha, 2015), but no recent integration of current theories and evidence has been published.

Some psychologists have been at the forefront of producing supportive research and theory; others have devoted much of their careers to criticizing the field. Both stances can be explained by psychologists’ expertise on relevant processes such as perception, memory, belief, and conscious and nonconscious processes. However, many psychologists probably lack solid knowledge of the area. An informed psi skeptic wrote, “Most psychologists could reasonably be described as uninformed skeptics—a minority could reasonably be described as prejudiced bigots—where the paranormal is concerned” (French, 2001, p. 7). It is thus important to have an overview and discussion of the research and theory on the topic. This article will (a) introduce the domain of psi research; (b) discuss relevant theoretical frameworks from physics, psychology, and evolutionary theory; (c) review recent/updated meta-analyses in the field; and (d) provide guidelines for future research.

The Domain of Psi Research

From the founding in 1882 of the Society for Psychical Research, research on psi has used or even developed sci-
entific practices, with the aim to “examine without prejudice or prepossession” the nature of these phenomena. Parapsychology can be defined as the study of purported psi phenomena using the scientific method, and the Parapsychological Association, the professional association of the field, has been an affiliate of the American Association for the Advancement of Sciences (the world’s largest general scientific society) since 1969.

Psi typically includes two major areas: (1) what used to be called extrasensory perception, or ESP, and (2) psychokinesis, or PK. ESP includes purported telepathy (being affected by someone’s thoughts or emotions, unmediated by the senses or logical inference, such as guessing more accurately than would be expected by chance who sends you an e-mail unexpectedly), clairvoyance (obtaining information about a distant state of affairs, unmediated by the senses or logical inference, such as in remote viewing (RV) in which someone accurately describes details of a place chosen at random by someone else), precognition/presentiment (being affected by an event taking a place in the future that could not have been foreseen, as in dreaming about planes crashing against tall buildings the night before 9/11), and retrocognition (having noninferable knowledge about a past event). ESP is a misleading term because it suggests perception as the mediating mechanism, although few if any psi researchers nowadays assume this to be the case. Furthermore, the distinction among these phenomena is a function of how they are tested or considered rather than of different mechanisms. Examples labeled as clairvoyance could also be considered as telepathy, and both of them could be subsumed under precognition, because someone at some point in the future will find out that information. The term anomalous cognition will be used in this article (Cardenä et al., 2015; May, Utts, & Spottiswoode, 1995).

PK refers to putative direct action of mental events (e.g., intention) on physical objects, unmediated by muscular or indirect mechanical activity. There is macropsychokinesis (or anomalous force), an effect on observable objects such as a table levitating without any apparent mechanical explanation, and micropsychokinesis (or anomalous perturbation), an effect on small, unobservable events, such as mentally affecting the output of a random number generator that otherwise produces random outputs. Some psi researchers study the possibility of consciousness surviving death, including studies of children who spontaneously report information about a past life to which neither they or those close to them apparently had access (Mills & Tucker, 2015), but which have also been interpreted as examples of anomalous cognition rather than of survival (Sudduth, 2009). Both descriptive and experimental approaches can be employed to evaluate psi phenomena.

At its inception, psychology and parapsychology were not clearly distinct disciplines, and foundational figures of the former also supported the latter (Cardenä, 2015a; Sommer, 2013. They include Bekhterev, Hans Berger (inventor of the electroencephalogram), Binet, Fechner, Sigmund Freud, Luria, Ramón y Cajal, and American Psychological Association (APA) presidents William James and Gardner Murphy. More recently, faculty from top-ranked universities such as Harvard, Princeton, and Stanford, including a past APA president, endorsed continuing research on psi (Cardenä, 2014).

Parapsychology has also contributed to methods and subject areas later integrated into psychology, among them the first use of randomization along with systematic use of masking procedures (Hacking, 1988); the first comprehensive use of meta-analysis, in 1940 (Gupta & Agrawal, 2012); study preregistration since 1976 (Johnson, 1976); and pioneering contributions to the psychology of hallucinations, eyewitness reports, and dissociative and hypnotic phenomena (for a review, see Hövelmann, 2015).

Psi Phenomena and Physics Theories

Because psi phenomena are sometimes assumed a priori to violate physics principles, three common objections will be discussed, namely that (1) they violate the “laws of nature”; (2) if accepted, they would invalidate scientific achievements; and (3) there are no theories to account for them. What is often meant by psi critics as violations of the laws of nature involves assumptions about an event not being able to affect another at a distance without some mediating form of known energy, future events being unable to affect previous ones, and mental events not having direct effects on other than the organism privy to them. However, quantum mechanics (QM) and Einstein’s theory

Etzel Cardeña

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
of relativity have depicted a reality that differs substantially from commonsensical assumptions. Nobel laureate and pioneer of molecular biology Max Delbrück (1986) expressed it so: “Modern science . . . has forced us to abandon absolute space and time, determinism, and the absolute object” (p. 279).

Nonlocality

In his interpretation of QM (and experts differ on how to interpret it, e.g., Schlosshauer, Kofler, & Zeilinger, 2013), the eminent physicist Bernard d’Espagnat (1979, 2006) discussed the implications of experiments showing that measuring/observing the property of a particle, such as its spin, instantaneously determines that of another particle entangled with it, no matter how distant. Entanglement means that the quantum states of such particles are not independent but part of a system, which can be produced in different ways. D’Espagnat concluded that such experiments falsify the local realist theory that effects cannot propagate faster than light and that objects far apart in space are relatively independent. For him, the world is not made of separate “material” objects embedded in space-time, but of a nonseparable, indivisible field, a “veiled reality,” with which consciousness interacts. He concluded that the implications of QM and “transcendentalism-inclined thinkers” (d’Espagnat, 2006, p. 429) have points in common, as did renowned physicist David Bohm (1986) in his theory of the implicate order or guiding field, which he applied to psi phenomena.

Scientific American journalist George Musser (2015) also supported a nonlocal interpretation of QM and considered space “a doomed concept” (p. 125). He also described how effects violating assumptions of locality do not occur exclusively at the particle but also at the cosmic level (and at the mesolevel of living beings; see Lambert et al., 2013). Along these lines, Princeton physics philosopher Hans Halvorson concluded that a form of superentanglement links every aspect of everything in the universe (Musser, 2015, p. 139). In principle, thus, psi phenomena—such as a sudden death affecting a loved one at another location—are consistent with a nonlocal view of the universe. Furthermore, as compared with classical physics, which depicts a universe where everything is determined by previous causes, QM proposes that before there is a collapse of the quantum wave function by some type of measurement, objects are only probability functions (Musser, 2015). In parapsychology, observational theories propose that psi experiments exploit the indeterminacy of a system, which may become slightly biased by the intention of an observer (Millar, 2015), or as Stapp (2017) put it, by “relevant conditions that include the experienced emotions of biological agents” (p. 106).

Time

Einstein’s theories of time and the ensuing experiments demonstrated that objectively measured time and space are not absolute and depend on such variables as the position and speed of the observers and the gravitational field. For instance, events that are still in the future of a slow-moving individual may have already occurred to a faster moving one; furthermore, in the special relativity block universe theory of time, past, present, and future coexist simultaneously although we experience only the present (Davies, 2002). Despite our typical perception of time as only an ever-receding series of moments, experiments on quantum retrocausality (or backward causation) suggest that future events may affect previous ones. For example, measuring the spin of a particle, which collapses its probabilistic wave function into a determinate value, seems to retroactively determine the spin of a delayed photon entangled with it. Physicist Daniel Sheehan (2011, 2015) concluded that experiments in physics and psi support retrocognitive effects, and physicist and parapsychologist Edwin May has developed a theory in which signals from a future space-time point, such as having eventual knowledge of the target of an experiment, may affect previous cortical processes of those trying to guess it (e.g., Marwaha & May, 2016). An alternative explanation is that consciousness may bias a future event (Stapp, 2017).

Consciousness/Sentience

One of the interpretations of QM requires that the measurement that makes a wave function of probabilities collapse into a determinate outcome be made by a sentient observer (Stapp, 2017). Consistent with a causal role for sentience, Delbrück (1986) criticized “the Cartesian cut between mind and matter” (p. 279), and cognitive psychologist Max Velmans (2000) also discussed the reasons why a hard distinction between “objective” and “subjective” phenomena is misguided. Along these lines, a professor of cosmology wrote that “the materialist position in physics appears to rest on shaky metaphysical ground” (Frank, 2017, quoted from the subtitle) and questioned the materialist stance in the neurosciences to explain consciousness. Renowned philosopher of mind Thomas Nagel (2012) concluded that the explanatory gap between neurochemical processes and mental experiences is difficult to resolve from a materialist, evolutionary perspective and that reality is not reducible to material, mental, or functional realms, but subsumes them all. Princeton physicist Freeman Dyson (1988, p. 297) ascribed different levels of mind from the particle to the cosmic levels, and Velmans (2000) concluded that a continuous model of sentience is more parsimonious than one proposing that mentality just emerges from matter at some level of complexity. To add eminent neuroscientists to those who endorse nonmaterialist views of mind, Christof
Koch, an earlier collaborator of arch-reductionist Francis Crick (Crick & Koch, 1990), concluded with Giulio Tononi that consciousness is a fundamental property of information in complex entities (Tononi & Koch, 2015, see also Dyson, 1988; Kelly, 2015).

But how might “mental” events interact with “physical” ones, assuming that they differ ontologically? University of London professor of mathematics and cosmology Bernard Carr (2015) has described recent hyperspatial or hyperdimensional approaches that posit additional dimensions beyond the temporal and three-spatial ones. He proposes that events that seem to be distant in our three-dimensional space may be adjacent in a hyperdimensional one, and that the dichotomy between mind and matter of common sense is resolved by a hyperdimensional “transcendental field” in which mental phenomena can have causal effects. Lawrence Livermore National Laboratory physicist Henry P. Stapp (2017, p. 65) has developed a “realistically construed orthodox quantum mechanics” model in which conscious intentions can produce a small bias on quantum processes, and indeed research has shown significant small effects of intention on photon wave patterns (Radin, Michel, & De- lorme, 2016).

The above views do not “prove” that psi phenomena exist but makes them plausible, and some physicists have proposed specific theories for them. They are also a response to psychologists who state that they psi phenomena are impossible (“Parapsychologists believe in ‘impossible’ things,” Alcock, 2010, p. 29; “(psi) conflicts with what we know to be true about the world,” Wagenmakers, Wetzels, Borsboom, & Van der Maas, 2011, p. 46).

Psychological and Evolutionary Theories of Psi

Two psychological theories, psi-mediated instrumental response (PMIR; Stanford, 2015) and first-sight theory (Carpenter, 2012), seek to integrate psi with psychological and evolutionary theories. Although varying in details, both propose that psi information continuously, although usually nonconsciously, impinges on mental processes and may serve adaptive and/or personal inclinations. PMIR has been the basis for studies in which rewarding psi tasks embedded within nonpsi experiments were found to affect performance in experiments (see section on implicit anomalous cognition below). A premise of PMIR is that an organism may respond to events outside of its sensory reach if it would respond to them if they were perceivable, such as avoiding an unperceivable dangerous situation, and there is a motivational component to what the organism will likely attend to depending on its particular dispositions and schemata (Stanford, 2015). Similarly, according to the first-sight model: (a) psi is not limited by the commonsensical view of time and space and is fundamental to all organisms, and (b) it mostly operates nonconsciously but may affect conscious-ness and action in accordance with the organism’s dispositions (Carpenter, 2012).

There are also explanations of why alterations of consciousness have been found to relate to psi. According to the “noise reduction” theory, psi information is subtle and likely to remain nonconscious in the midst of the overwhelming information provided by the senses and bodily actions unless these inputs are reduced (Honorton, 1977). Thus, procedures that reduce these stimuli—such as meditation, hypnosis, and ganzfeld—should facilitate awareness of psi (see the sections on ganzfeld and dream research below). Besides restriction of sensory input, alterations in consciousness may make awareness of psi more likely by reducing critical thought and stimulating a sense of interconnectedness (Cardeña, 2010).

Psi has also been discussed from an evolutionary perspective. According to Broughton (2015), psi should be seen in the larger context of biological processes including brain functioning and evolution. He stated that psi is a correlation between future and previous events that could have been maintained even providing as little as a 1% fitness advantage. This small effect could subtly affect decisions through hunches and similar mechanisms, consistent with the small effects found in research. Research and observations in nature support the existence of psi in other species (Safina, 2015; Sheldrake, 2015).

As to statements such as the one by a cognitive scientist that accepting psi phenomena would “send all of science as we know it crashing to the ground” (Hofstadter, 2011, para. 9) the most sensible answer is that psi phenomena are compatible with some interpretations by eminent physicists and manifest small effects that in no way invalidate the accomplishments of current science (Stapp, 2017). The Office of Technology Assessment (1989) report concluded that it is important to find out how psi can obtain “a fairer hearing across a broader spectrum of the scientific community, so that emotionality does not impede objective assessment of experimental results” (p. 337).

Summary of Meta-Analyses on Psi

This section summarizes recent or updated comprehensive meta-analyses of psi research found through a recent comprehensive anthology that reviewed meta-analyses in the field (Cardeña et al., 2015), contacts with parapsychology researchers, and an additional literature search. The latter used two databases, PsycInfo and Medline, without language or year restriction, using metaanalysis OR meta-analysis OR meta analysis as keywords, and the following as subject words: parapsychology, psi, telepathy, clairvoyance, precognition, psychokinesis, PK, anomalous healing, and intercessory prayer. The search produced about 20 nonoverlapping items, many of them not meta-analyses but comments about them. All comprehensive recent/updated
meta-analyses found are discussed below along with more limited but relevant ones. Although meta-analyses have limitations that may affect their results (e.g., a potential publication bias), they contribute to knowledge of established and contentious areas (Chan & Arvey, 2012). The primary sources in the meta-analyses reviewed here include alternative and comprehensive analyses and evaluate variables that might have impacted the data, including the design quality and homogeneity of the studies, and potential publication biases. Tables 1 and 2 include statistics for complete and homogenized (sometimes after trimming 10% of extreme results in heterogeneous data sets) meta-analyses when available.

### Anomalous Cognition

In anomalous cognition research, participants “guess” a randomly chosen target from a known (e.g., a set of cards) or unknown (e.g., a film clip from a large sample) set without information from the senses or logical inference. It includes two main models of research: free response and forced choice (for a review of safeguards commonly used to avoid confounds such as sensory leakage and judging bias see Palmer, 2015a).

#### Free response.

In free-response studies, the target that the masked raters will evaluate is not part of a set known to them such as a deck of cards, but of a large or undetermined dataset, such as a photo or film clip from an unknown large or open set (e.g., the stimuli may come from a pool of dozens of clips including animated, documentary, or feature films, or a location chosen at random). Various types of free-response protocols have been studied.

**Ganzfeld** is a German term for “whole field.” In psi research, it refers to a procedure in which the participant sits in a comfortable chair and listens to physical relaxation

<table>
<thead>
<tr>
<th>Database</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>$ES$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ganzfeld (adapted from Storm et al., 2010b)</td>
<td>108</td>
<td>8.31</td>
<td>$&lt;10^{-16}$</td>
<td>.142</td>
<td>—</td>
</tr>
<tr>
<td>Ganzfeld (hom)</td>
<td>102</td>
<td>8.13</td>
<td>$&lt;10^{-16}$</td>
<td>.135</td>
<td>[.10, .17]</td>
</tr>
<tr>
<td>New (all)</td>
<td>30</td>
<td>6.34</td>
<td>$1.15 \times 10^{-10}$</td>
<td>.152</td>
<td>—</td>
</tr>
<tr>
<td>New (hom)</td>
<td>29</td>
<td>5.48</td>
<td>$2.13 \times 10^{-8}$</td>
<td>.142</td>
<td>[.07, .22]</td>
</tr>
</tbody>
</table>

#### Precognition/Bem-type studies (based on Bem et al., 2015)

<table>
<thead>
<tr>
<th>Database</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>$ES$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bem et al. (all)</td>
<td>90</td>
<td>6.40</td>
<td>$1.2 \times 10^{-10}$</td>
<td>.09a</td>
<td>[.06, .11]</td>
</tr>
<tr>
<td>Bem et al. (fast)</td>
<td>61</td>
<td>7.11</td>
<td>$5.8 \times 10^{-13}$</td>
<td>.11</td>
<td>[.04, .14]</td>
</tr>
<tr>
<td>Bem et al. (slow)</td>
<td>29</td>
<td>1.38</td>
<td>.16</td>
<td>.03</td>
<td>[—, .00]</td>
</tr>
</tbody>
</table>

#### Psi dream studies (adapted from Storm et al., 2017)

<table>
<thead>
<tr>
<th>Database</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>$ES$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined (all)</td>
<td>52</td>
<td>5.01</td>
<td>$2.72 \times 10^{-7}$</td>
<td>.18</td>
<td>—</td>
</tr>
<tr>
<td>Combined (hom)</td>
<td>50</td>
<td>5.32</td>
<td>$5.19 \times 10^{-8}$</td>
<td>.20</td>
<td>[.11, .29]</td>
</tr>
</tbody>
</table>

#### Remote viewing (adapted from Storm, Tressoldi, and Di Risio, 2015; Dunne & Jahn, 2003; Milton, 1997)

<table>
<thead>
<tr>
<th>Database</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>$ES$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRI (770)</td>
<td>(770)</td>
<td>.20</td>
<td>[—, .23]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAIC (445)</td>
<td>.23</td>
<td>[.19, .27]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milton (hom)</td>
<td>75 (2,682)</td>
<td>5.85</td>
<td>$2.46 \times 10^{-9}$</td>
<td>.17</td>
<td>[.10, .22]</td>
</tr>
<tr>
<td>Dunne &amp; Jahn (653)</td>
<td>.21</td>
<td>[.18, .24]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bierman &amp; Rabeyron (550)</td>
<td>.27</td>
<td>[.23, .31]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994–2014 (314)</td>
<td>.39</td>
<td>[.14, .64]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Presentiment (adapted from Mossbridge, Tressoldi, & Utts, 2012)

<table>
<thead>
<tr>
<th>Database</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>$ES$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mossbridge et al. (all)</td>
<td>26</td>
<td>5.3</td>
<td>$5.7 \times 10^{-8}$</td>
<td>.21</td>
<td>[.13, .29]</td>
</tr>
<tr>
<td>Mossbridge et al. (hq)</td>
<td>13</td>
<td>4.4</td>
<td>$6 \times 10^{-6}$</td>
<td>.24</td>
<td>[.13, .35]</td>
</tr>
<tr>
<td>Mossbridge et al. (lq)</td>
<td>13</td>
<td>2.96</td>
<td>$&lt;.002$</td>
<td>.17</td>
<td>[.06, .29]</td>
</tr>
</tbody>
</table>

#### Forced choice (adapted from Baptista et al., 2015)

<table>
<thead>
<tr>
<th>Database</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>$ES$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honorton/Ferrari (all)</td>
<td>309</td>
<td>11.41</td>
<td>$6.3 \times 10^{-25}$</td>
<td>.020</td>
<td>[.09, .31]</td>
</tr>
<tr>
<td>Honorton/Ferrari (hom)</td>
<td>248</td>
<td>6.02</td>
<td>$1.1 \times 10^{-9}$</td>
<td>.012</td>
<td>[.05, .19]</td>
</tr>
<tr>
<td>STDR (all)</td>
<td>91</td>
<td>10.82</td>
<td>$10^{-16b}$</td>
<td>.04</td>
<td></td>
</tr>
<tr>
<td>STDR (hom)</td>
<td>72</td>
<td>4.36</td>
<td>$6.5 \times 10^{-6b}$</td>
<td>.01</td>
<td>[.01, .02]</td>
</tr>
</tbody>
</table>

**Note.** $k$ = number of studies; $Z$ = cumulative standard deviation from the mean; $ES$ = mean effect size; CI = confidence interval for $ES$; fast = protocols involving fast-thinking processes; slow = protocols involving slow-thinking processes; hom = homogeneous; SRI = Stanford Research Institute; SAIC = Science Applications International Corporation; hq = high-quality study subset; lq = low-quality study subset; STDR = Storm, Tressoldi, and Di Risio.

* $P$-curve analysis = .20.  
  * One-tailed, quality weighed.
but had they used the (apparently indicated) exact binomial of a comprehensive previous database (Storm & Ertel, 2001) for all studies, along with a psi effect. An earlier and more limited meta-analysis by Storm et al. (2010b) and Williams (2011) supported most recent and comprehensive meta-analyses of the data-

Milton and Wiseman (1999) did not find a significant effect, and is the most consistently supportive database for psi of the last few decades. The methodological development of ganzfeld research followed a joint communiqué by psi-critic Ray Hyman and psi-proponent Charles Honorton (Hyman & Honorton, 1986) on how to conduct the experiments. The psi ganzfeld technique is based on the “noise reduction” theory mentioned above.

Research on ganzfeld has been meta-analyzed repeatedly and is the most consistently supportive database for psi of the last few decades. The methodological development of ganzfeld research followed a joint communiqué by psi-critic Ray Hyman and psi-proponent Charles Honorton (Hyman & Honorton, 1986) on how to conduct the experiments. The most recent and comprehensive meta-analyses of the database by Storm et al. (2010b) and Williams (2011) supported a psi effect. An earlier and more limited meta-analysis by Milton and Wiseman (1999) did not find a significant effect, but had they used the (apparently indicated) exact binomial test, it would have (Storm et al., 2010b, p. 473).

Table 1 shows the meta-analyses for (a) the aggregation of a comprehensive previous database (Storm & Ertel, 2001) with the newer database and for (b) the newer database alone (Storm et al., 2010b) for all studies, along with analyses for (c) the aggregation of all homogeneous combined studies and for (d) the homogeneous newer database, for which an outlier with a very high supportive z score was excluded. For the latter, selected participants (based on previous experience with the protocol and/or traits associated with psi performance such as being a meditator) had a bigger effect size, $ES = 0.26$, than their counterparts, $ES = 0.05$. Williams (2011) reported that in ganzfeld participants guess around 31% of the time the correct target out of four choices presented in random order, when mean chance expectation would be 25%.

Hyman (2010) criticized the Storm et al. meta-analysis, claiming that meta-analyses should be conducted prospectively and that psi cannot be shown on demand. He also described a ganzfeld study that did not replicate the effect, although he disregarded others that did. Storm et al. (2010a) responded that other accepted phenomena in science cannot be produced on demand, and that retrospective meta-analyses are routinely used in science.

Rouder, Morey, and Province (2013) conducted a Bayesian probability analysis of the newer Storm et al. dataset excluding studies that had used manual instead of computerized randomization. They concluded that the Bayesian factor decreased from 6 billion to 1 to circa 330 to 1, but that because of a lack of a plausible mechanism and the possibility of unpublished replication failures, the meta-analysis did not support psi. They added, however, that the degree of evidence was “greater than that provided in many routine studies in cognition” (p. 245). Storm, Tressoldi, and Utts (2013) conducted a Bayesian analysis of a binomial distribution.

**Table 2**

**Summary of Meta-Analytic Findings for Anomalous Perturbation**

<table>
<thead>
<tr>
<th>Database</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote influence</td>
<td>(adapted from Schmidt, 2015)</td>
<td>36</td>
<td>.001</td>
<td>.106</td>
<td>[0.04, 0.17]</td>
</tr>
<tr>
<td>EDA-DMILS (hom)</td>
<td>15</td>
<td>.013</td>
<td>.128</td>
<td></td>
<td>[0.03, 0.23]</td>
</tr>
<tr>
<td>Attention facilitation</td>
<td>11</td>
<td>.029</td>
<td>.114</td>
<td></td>
<td>[0.01, 0.22]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Noncontact healing studies (adapted from Roe, Sonnex, &amp; Roxburgh, 2015)</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonhuman (all)</td>
<td>49</td>
<td>&lt;.05</td>
<td>.258</td>
<td></td>
<td>[0.24, 0.28]</td>
</tr>
<tr>
<td>Nonhuman (hq)</td>
<td>22</td>
<td>&lt;.05</td>
<td>.115</td>
<td></td>
<td>[0.09, 0.14]</td>
</tr>
<tr>
<td>Human (all)</td>
<td>57</td>
<td>&lt;.05</td>
<td>.203</td>
<td></td>
<td>[0.18, 0.23]</td>
</tr>
<tr>
<td>Human (hq)</td>
<td>27</td>
<td>&lt;.05</td>
<td>.224</td>
<td></td>
<td>[0.19, 0.25]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dice (adapted from Radin &amp; Ferrari, 1991)</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radin &amp; Ferrari (all)</td>
<td>73</td>
<td>18.2</td>
<td>&lt;.001</td>
<td>.0072</td>
<td>[.0065, 0.0079]</td>
</tr>
<tr>
<td>Radin &amp; Ferrari (hom)</td>
<td>59</td>
<td>3.19</td>
<td>&lt;.001</td>
<td>.0029</td>
<td>[.0017, 0.0041]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Micro-PK (adapted from Bösch, Steinkamp, &amp; Boller, 2006)</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNG (all)</td>
<td>380</td>
<td>2.47</td>
<td>&lt;.05</td>
<td>.50003</td>
<td>[.0017, 0.0041]</td>
</tr>
<tr>
<td>RNG (–3)</td>
<td>377</td>
<td>4.08</td>
<td>&lt;.001</td>
<td>.50028</td>
<td>[.0017, 0.0041]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Global Consciousness Project (GCP; Nelson, 2015, personal communication, 2016)</th>
<th>$k$ (trials)</th>
<th>$Z$</th>
<th>$p$</th>
<th>ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCP</td>
<td>461</td>
<td>7.23</td>
<td>2.34 × 10$^{-13}$</td>
<td>.33</td>
<td></td>
</tr>
</tbody>
</table>

Note. $k$ = number of studies; $Z =$ cumulative standard deviation from the mean; $ES =$ effect size; $CI =$ confidence interval for $ES$; EDA = electrodermal activity; DMILS = direct mental interaction in living systems; hom = homogeneous; hq = high-quality study subset; RNG = random number generator.

$^a$ ES weighed by methodological quality. $^b$ Mean $\pi$ for a binomial distribution.

instructions and exposure to white or pink noise (unpatterned random frequencies, similar to the sound between radio stations), with two acetate ovals covering the eyes in front of which red light bulbs produce the effect of shapeless redness. The participant’s task is to become aware of an unknown image or clip chosen randomly, which might be shown simultaneously in a distant computer with nobody watching it (clairvoyance), someone watching it (telepathy), or is chosen after the participant makes a selection (precognition). The psi ganzfeld technique is based on the “noise reduction” theory mentioned above.

Research on ganzfeld has been meta-analyzed repeatedly and is the most consistently supportive database for psi of the last few decades. The methodological development of ganzfeld research followed a joint communiqué by psi-critic Ray Hyman and psi-proponent Charles Honorton (Hyman & Honorton, 1986) on how to conduct the experiments. The most recent and comprehensive meta-analyses of the database by Storm et al. (2010b) and Williams (2011) supported a psi effect. An earlier and more limited meta-analysis by Milton and Wiseman (1999) did not find a significant effect, but had they used the (apparently indicated) exact binomial test, it would have (Storm et al., 2010b, p. 473).

Table 1 shows the meta-analyses for (a) the aggregation of a comprehensive previous database (Storm & Ertel, 2001) with the newer database and for (b) the newer database alone (Storm et al., 2010b) for all studies, along with analyses for (c) the aggregation of all homogeneous combined studies and for (d) the homogeneous newer database, for which an outlier with a very high supportive z score was excluded. For the latter, selected participants (based on previous experience with the protocol and/or traits associated with psi performance such as being a meditator) had a bigger effect size, $ES = 0.26$, than their counterparts, $ES = 0.05$. Williams (2011) reported that in ganzfeld participants guess around 31% of the time the correct target out of four choices presented in random order, when mean chance expectation would be 25%.

Hyman (2010) criticized the Storm et al. meta-analysis, claiming that meta-analyses should be conducted prospectively and that psi cannot be shown on demand. He also described a ganzfeld study that did not replicate the effect, although he disregarded others that did. Storm et al. (2010a) responded that other accepted phenomena in science cannot be produced on demand, and that retrospective meta-analyses are routinely used in science.

Rouder, Morey, and Province (2013) conducted a Bayesian probability analysis of the newer Storm et al. dataset excluding studies that had used manual instead of computerized randomization. They concluded that the Bayesian factor decreased from 6 billion to 1 to circa 330 to 1, but that because of a lack of a plausible mechanism and the possibility of unpublished replication failures, the meta-analysis did not support psi. They added, however, that the degree of evidence was “greater than that provided in many routine studies in cognition” (p. 245). Storm, Tressoldi, and Utts (2013) conducted a Bayesian analysis of a binomial distribution.
done, and concluded that their results did support a psi effect. Baptista, Derakhshani, and Tressoldi (2015) conducted additional analyses on the ganzfeld data that ratified a psi effect and rectified general claims that (a) when ganzfeld study quality goes up, ES goes down (actually, the opposite seemed to be the case), (b) ES had decreased in more recent studies (it has not), (c) psi generally declines in the course of a long study (it does not), and (d) a file-drawer analysis of a reasonable number of unreported nonsignificant results would annul the significant results (it does not).

In implicit anomalous cognition studies, volunteers respond to a psychological task, with a hidden psi aspect to it. As an example, in one study, participants were part of a research dyad and one of them had to indicate esthetic preference for Kanji Japanese characters. Unbeknownst to them, they were being tested for a psi target selected randomly. When participants chose the psi character target, their research partners did a more pleasant task than those whose partners did not select the psi target (Watt & Nagtegaal, 2000). The outcome variable was whether participants chose the target more often than would be expected by chance. Although there has not been a meta-analysis of these studies, Palmer (2015b, p. 227) concluded in a review that studies with a hidden reward had more significant outcomes than would be expected by chance.

Related to this paradigm, studies designed by Cornell psychologist Daryl Bem (2011) tested the hypothesis that a future stimulus might have a retroactive influence on a previous response. Bem took mainstream priming studies, in which a preceding word or image affects an ensuing response, and “time-reversed” them so that the word or image is presented after the response of the volunteer. For instance, one of his tasks evaluated whether a valence-consistent or inconsistent word affects the response time of a preceding image. Bem (2011) reported on nine different protocols with more than 1,000 participants and found that all but one of them was independently significant and that the mean ES was significant (the analyses were one-tailed, but they would have remained significant with two-tailed tests). Thus, the results supported the interpretation that a stimulus occurring later may influence a previous response more often than would be expected by chance.

Bem’s studies, published in a major psychology journal, caused a storm of commentaries (Cardeña, 2015b). A meta-analysis of two of Bem’s experiments, along with attempted replications by the authors of the article and by other investigators, concluded that a Bayesian analysis showed no psi effect (Galak, LeBoeuf, Nelson, & Simmons, 2012). Table 1 shows a larger a meta-analysis of all replication attempts until then, 90 experiments from 33 laboratories at the time of publication (Bem, Tressoldi, Rabeyron, & Duggan, 2015). The overall effect was significant (as was a Bayesian analysis), and the ES for the complete database and the independent replications (excluding Bem’s experiments, P-curve analysis = 0.24) were similar. The authors report that hundreds of unpublished experiments with low ES would be required to annul the significant results of their meta-analysis. The authors also classified the replications into two groups: five protocols involving automatic, “fast-thinking” unconscious processing and two protocols involving “slow-thinking,” deliberative processing (cf. Kahneman, 2011). All the significant results belonged to the “fast-thinking” group and the most successful one used erotic stimuli, in general agreement with the theories reviewed earlier that posit psi as a mostly nonconscious process geared to future reinforcers. The two “slow-thinking,” deliberative protocols were not singly or jointly significant. Table 1 shows the cumulative results divided by categories.

In everyday life, ostensible anomalous cognition often occurs during dreams (Kelly & Tucker, 2015). The first comprehensive analysis of controlled studies was carried out by Yale psychologist Irvin Child (1985) on the dream psi studies conducted at the Maimonides Medical Center sleep lab. This protocol involved waking up (usually selected) participants after they had been in a REM sleep stage, which is strongly associated to dreaming, and querying them for their dream content. The task of the participant was to dream about an unknown image chosen at random by the researchers either while participants slept or at a later time. Child reported that in 20 out of 25 experiments the dream content on average had been correctly matched (blindly) to the target directly or on the top half of a binary division of multiple choices at a better than chance level, with a probability against chance of $1.46 \times 10^{-8}$. Radin (2006) estimated that the target had been judged to be on the top half of the distribution a highly significant 63% of the times (50% being mean chance expectation).

After the Maimonides program, most researchers have relied on dream diaries, which produce much poorer dream recollection than REM awakenings, rather than on studies in sleep labs, which are far more demanding. Storm et al. (2017) metaanalyzed the Maimonides and post-Maimonides studies. Table 1 includes the analysis for all and a homogeneous set, showing support of the psi hypothesis. They also report increased design rigor across time, and no association between study quality (rated blindly according to seven criteria including appropriate randomization, good masking, and so on) and ES. They also conducted a Bayesian analysis that confirmed their results. Their meta-analysis includes a large, well-controlled study by Watt (2014) in which independent, masked judges matched at better than chance level dream reports to the film clip that participants later saw. When alerted to a potential effect of dropouts from the study, Watt reanalyzed her data and still found a significant effect for her planned test ($p = .04$; Watt & Valášek, 2015).
RV (remote viewing) is a technique in which an individual describes a place, chosen at random, where a sender is located at the present or at a future time (there may also be just a location chosen without any observer there). Afterward the description is used to select the target among different possibilities. Associative RV is a type of precognitive RV in which the participant tries to guess a target to be selected in the future, and which may be associated with a particular event, for instance a change in the stock market. 

Table 1 shows Baptista et al.’s (2015) summary of the available data (the dataset for Milton, 1997, is homogeneous after deleting three studies). The confidence intervals of the data sets are of a similar magnitude and do not include 0.00, which would indicate no effect. The analyst for the first two data sets (Stanford Research Institute and Science Applications International Corporation; Utts, 1996) wrote that RV volunteers who had participated in previous research exhibited a greater ES (0.38) than novices (0.16). The psi skeptic Hyman (1995) concluded that the Science Applications International Corporation experiments were “well-designed and the investigators have taken pains to eliminate the known weaknesses in previous parapsychological research . . . I cannot provide suitable candidates for what flaws, if any, might be present.” Nonetheless, objections have been raised to the Dunne and Jahn (2003) database for sampling without replacement and not always selecting the targets randomly, but even if that dataset is eliminated the overall effect remains significant (Baptista et al., 2015).

In presentiment research, physiological activity preceding an unpredictable stimulus is hypothesized to anticipate the response that follows it, for instance that the preceding skin conductance to emotionally-charged stimuli will differ from that of neutral stimuli, interspersed randomly. Mossbridge, Tressoldi, and Utts (2012) meta-analyzed relevant studies published between 1978 and 2010. Table 1 shows that the overall ES for a physiological response preceding the stimulus was significant. The authors also conducted a masked preanalysis evaluation of the quality of the studies (according to level of peer review, type of random number generator, and whether an analysis of expectation bias had been conducted), and arrived at a division of 13 higher and 13 lower quality sets. Although both ESs were significant, the higher quality studies had a bigger ES than the lower quality ones (see Table 1). Mossbridge and collaborators (2015) later responded point-by-point to the criticisms of their meta-analysis raised by Schwarzkopf (2014).

Forced choice. In forced-choice studies, the guessing possibilities are finite and the possibilities are known by the person, for instance cards in a randomized deck. The protocol measures whether the participant can guess correctly more often than would be expected by chance. This was a common design in the middle of the 20th century. Honorton and Ferrari (1989) conducted a meta-analysis of forced-choice precognition research conducted between 1935 and 1987 by 62 investigators. Table 1 shows the analyses for all of 309 experiments and for the 248 homogeneous ones, revealing highly significant but very small ESs. They also reported that the ES had remained constant through the decades, that there was no relation between study outcome and an index of design quality based on eight criteria (e.g., preplanned analysis and randomization method), and that a file-drawer effect could not reasonably explain away the results. They also identified the 17 best studies, with selected samples and trial-by-trial feedback, which produced the largest effect of any other groups of studies in their database, \( Z = 15.84, ES = 0.12 \). The results of selected participants (based on prior performance) were better than those of their counterparts, \( t(246) = 3.16, p = .001 \) (Honorton & Ferrari, 1989, see also Baptista et al., 2015).

A second meta-analysis of forced-choice experiments was carried out by Storm, Tressoldi, and Di Risio (2012) on 91 studies conducted between 1987 and 2010, and on 72 homogeneous studies (see Table 1). They concluded that there was a small but significant effect, and no evidence that the results could be explained by low-quality designs (based on six criteria including appropriate randomization and random target positioning) or selective reporting, and that ESs had increased across time. Baptista et al. (2015) reported that the mean ES of the studies with selected participants was larger than that of unselected ones (ES = .05 vs. ES = .008, \( p = .001 \)).

**Anomalous Perturbation**

Anomalous perturbation refers to the ostensible influence of intention on nonobservable systems, evaluated statistically (there are no meta-analyses of anomalous force). Remote influence research evaluates the effect of intentional efforts to change a parameter in a distant living system, unmediated by known physical means. Schmidt (2015) summarized his meta-analyses of three areas: (a) direct mental interaction in living systems, such as measuring the electrodermal activity (EDA) of a receiver while a distant agent is, at random times, trying to make that person aroused or calm; (b) remote staring, or changes in the EDA of a receiver as an agent looks at him/her through video at random times from a separate room; and (c) remote helping (or attention-focusing facilitation), in which a remote helper tries at random times to help a meditator focus on a target. Table 2 shows that all three research paradigms were supportive of psi. Schmidt (2015) wrote that the similar ES for the three areas validate each other and suggest the same underlying phenomenon. Nonetheless, there were some differences. The EDA–direct mental interaction in living systems dataset did not include four studies with inadequate randomization, and in the remaining 36 homogeneous studies dataset, there was a negative correlation \( (r = -.40) \) between ES and quality of study (based on 17 items includ-
produced a highly significant but small effect in the ex-
affect the fall of dice without touching them, and which
than 2 million dice throws, in which participants intended to
results should be taken cautiously.
Ferrari (1991) meta-analyzed 148 studies involving more
machine to avoid possible manipulation, was a common
variables as treatment allocation randomization and good masking pro-
Because there were negative correlations between indices of experimental quality and ES, the authors selected
the 22 studies rated as having good designs. Although the ES diminished, it remained significant (see Table 2). For
research on humans, the authors identified 57 studies with a physical bias of
ting adequate randomization and preregistration). Thus, studies were weighed according to sample size and quality
of the remote staring dataset was homogeneous
involving biological tissues or whole living beings other than the person having the intention. Two previous meta-analyses on intercessory prayer, partly overlapping, came to opposite conclusions as to whether there was a valid effect (Hodge, 2007; Masters, Spielmans, & Goodson, 2006). More comprehensively, Roe, Sonnex, and Roxburgh (2015) meta-analyzed two
types of studies: those relating to “nonwhole human stud-
animals, plants, and in vitro cultures) and those to “whole humans.” They pointed out that whereas in the
in the second category it would be difficult in some studies to discount the role of a placebo effect, the results for the first
category were unlikely to depend on such mechanisms as unconditioned expectancies.
For the nonwhole human studies, Roe et al. (2015) ident-
ted 49 heterogeneous studies, which had a significant
Weighting was carried out without awareness of results or authorship and included such variables as
treatment allocation randomization and good masking pro-
Because there were negative correlations between indices of experimental quality and ES, the authors selected
the 22 studies rated as having good designs. Although the ES diminished, it remained significant (see Table 2). For
research on humans, the authors identified 57 studies with adequate methodology, which produced a significant result (see Table 2), but because there was a negative correlation between design quality and ES, they analyzed the 27 studies with better methodology, and the ES remained significant (see Table 2). Roe et al. warned that because the funnel plot of the ESs suggest publication bias (and some authors did not provide exact probability values when not significant), results should be taken cautiously.

**Dice.** Trying to affect the fall of dice, typically in a machine to avoid possible manipulation, was a common
research paradigm used in the mid-20th century. Radin and Ferrari (1991) meta-analyzed 148 studies involving more
than 2 million dice throws, in which participants intended to affect the fall of dice without touching them, and which
produced a highly significant but small effect in the expected direction (Z = 19.68; \( p < .01 \)), with Z values decreasing and methodological quality improving over time. This dataset includes studies with a physical bias of
using higher dice faces as targets, so Table 2 shows significant but very small results for 73 studies after controlling
for this artifact, as well as for a homogenized set of 59 studies, indicating that the dice fell more often on the face
intended than would be expected by chance. For the homogenized dataset, ES did not relate to design quality. In
comparison, in the 31 control studies the selected face did not fall more often than would be expected by chance, \( Z = 0.36, p > .05 \).

**Micro-PK.** Bösch, Steinkamp, and Boller (2006) meta-
analyzed 380 studies on attempts to affect random number generators (RNGs). Table 2 shows significant but very
small effects for a random-effects model on a dataset in-
cluding and excluding the three largest studies. Although
the 137 control studies in which there was no intention to affect the RNGs did not show a significant deviation, \( Z = -1.51, p = .13 \), the authors concluded that the results might be explained by publication bias because there was
great heterogeneity and fewer studies below \( p < .05 \) and
\( p < .01 \) levels than would be expected by chance. In a later
reanalysis of the Bösch et al. data, Varvoglis and Bancel
(2015, p. 274) concurred that the distribution of significance
levels suggested some publication bias, but posited that an “extremely large” and unrealistic file drawer effect would be required to annul the results. They proposed instead that the data heterogeneity could be explained by the talent
(methodological and perhaps parapsychological) of partic-
ular experimenters and the far better than average perfor-
mance of two participants in the PEAR (Princeton Engi-
neering Anomalies Research) dataset, who contributed a
quarter of the data with zs of 5.6 and 3.4 as compared with 0.8 for the remaining participants.

In implicit anomalous perturbation, the experimenter sets hidden or secondary RNGs to be influenced by participants
without any necessary conscious intention. These studies partially inspired a research program, having been con-
ducted now for more than 15 years, known as The Global
Consciousness Project (Nelson, 2015). Its premise is that
events that simultaneously impact many people throughout
the world (e.g., the 9/11 attacks) create a coherence in
human consciousness that affects the randomness of a net-
work of 65 RNGs located in various countries. The collec-
tive RNG output from a time window around such major
events is compared with times in which no such events occur. Table 2 shows the analysis of 461 events, with a
significant result and a sizable ES at the level of the event
(Nelson, 2015, and personal communication, 2016). By
their nature, these data come from a single source, but data
and analyses are accessible at noosphere.princeton.edu/results .html#alldata.

Comparing all of these meta-analyses, there are consistent
patterns. First, overall the meta-analyses have been support-
ive of the psi hypothesis, with those that have not (e.g.,
Galak et al., 2012; Milton & Wiseman, 1999) generally
superseded by alternative, more comprehensive meta-
analyses. Second, the analyses relating to free-response
paradigms have the highest ESs, ranging from 0.11 to 0.39,
with most over 0.2. Then follow ESs for remote, noncontact
influence, ranging from 0.10 to 0.26. The ESs for forced-choice research, ranging from 0.01 to 0.04, are about one order of magnitude smaller and the anomalous perturbation ESs for dice and micro-PK (excluding the Global Consciousness Project) are also very small. Third, selected participants seem to evidence more psi than nonselected ones.

With respect to the last point, the most consistent data sets (ganzfeld, dream studies, and RV) have often used selected participants, and the analyses reviewed earlier for ganzfeld, forced choice, and micro-PK strongly support this practice (for research with “gifted” individuals under controlled conditions, see Edge, Morris, Palmer, & Rush, 1986). Characteristics shown to increase the likelihood of performing well in a psi experiment include a belief that one will do well in the study, some psychological traits (e.g., extraversion and openness to experience), a mental practice such as meditation, and previous experience in a psi experiment (for a review, see Cardeña & Marcusson-Clavertz, 2015). In a recent meta-analysis on forced-choice experiments, performance correlated positively with belief in psi, \( r = .13, p = .002 \), extraversion, \( r = .08, p = .02 \), and openness to experience, \( r = .12, p = .02 \) (Zdrenka & Wilson, 2017). Artists tend to score better than chance and other groups (Holt, Delanoy, & Roe, 2004).

There is evidence that testing while a participant is in a different state of consciousness than the ordinary, waking one is conducive to psi performance. Two of the more successful paradigms involve naturally occurring or induced alterations (i.e., ganzfeld, dreaming). In an earlier review, Honorton (1977) compared performance in psi studies involving hypnosis, meditation, induced relaxation, and ganzfeld and concluded that they produced better results than would be expected by chance, ranging from \( 1.2 \times 10^{-5} \) to \( 6 \times 10^{-12} \) (for more recent reviews, see Cardeña et al., 2015). Storm et al. (2010b, p. 476) compared research with ganzfeld, other purported psi-enhancing techniques such as meditation, and those not using psi-enhancing techniques. Ganzfeld had the largest ES (0.14), followed by other enhancing techniques (ES = 0.11) and studies without techniques (ES = −0.03), the last one differing from ganzfeld (ES mean difference = 0.17, \( p = .005 \)). The effect of ganzfeld may be mediated by how much it alters the state of consciousness (Marcusson-Clavertz & Cardeña, 2011; Roe, Hodrien, & Kirkwood, 2012).

**Discussion**

This overview of meta-analyses of various different research protocols supports the psi hypothesis. The analyses satisfy the “local and global criteria” specified by a critic of psi who demanded replicability, consistency of effects, and cumulativeness (Office of Technology Assessment, 1989). The meta-analyses, conducted on studies using different protocols and by different researchers, provide cumulative vertical and horizontal support of psi. Vertical in the sense that across time different protocols have continued to produce positive results beyond what would be expected by chance, and with increasing methodological rigor; horizontal in the sense that there is support for psi across research areas. If only one or a few protocols out of 10 were significant and the rest were not, it would be easier to speculate that the supportive results might be due to an artifact. In addition, the rigor of the psi meta-analyses has increased with time and typically include evaluation of possible selective reporting, quality of studies, and so on. The article will now consider some common criticisms of psi.

**If Psi Phenomena Are Real, Why Do Not All Studies Replicate Them?**

Considering the small ES found and potential sources of variability, including psychological and perhaps parapsychological experimenter effects (Palmer & Millar, 2015), one should expect some studies not to replicate (cf. Barrett, 2015; Lewontin, 1994). As Harvard professor Robert Rosenthal (1990) opined,

> Given the levels of statistical power at which we normally operate, we have no right to expect the proportion of significant results that we typically do expect, even if in nature there is a very real and very important effect. (p. 16; see also Utts, 1991)

Clearly, psi effects cannot be replicated “on demand,” but to put this fact in perspective, consider the “Many Labs” project, in which 36 independent laboratories attempted to replicate 16 psychology studies published in top journals, and only 34% of the replications fell within the confidence intervals of the original study (Open Science Collaboration, 2015). There were more consistent results when the replication was similar to the original study (Gilbert, King, Pettigrew, & Wilson, 2016), as has been found in ganzfeld (Bem, Palmer, & Broughton, 2001) and precognition (Bem et al., 2015) research.

**Why Are the Effect Sizes Typically So Small and How Do They Compare to Other Areas?**

As compared with real-life circumstances, psi experiments involve impersonal stimuli of little or no consequence, in contrast with reputed psi phenomena observed in everyday life (e.g., unexpected deaths of close people). Furthermore, psi seems to be more reliably manifested by only a few people, so the ESs are probably the average of larger effects of selected participants and smaller to null effects of others (cf. Harris & Rosenthal, 1988). As far as a comparison to other areas, in their analysis of more than 25,000 social psychology experiments, Richard, Bond, and Stokes-Zoota (2003) reported an average ES = 0.21, similar to some of the meta-analyses in Tables 1 and 2. The ES of
some psi protocols is not only comparable but much larger than those of the clinically recommended uses of aspirin or propanolol to prevent heart conditions (Spencer, 1995; Utts, 1991) and would be classified as “evidence-based” applying the criteria of clinical practice (cf. Haidich, 2010).

Aren’t the Significant Effects in Psi Produced by Low-Quality Experiments?

First, most of the meta-analyses reviewed controlled for quality and still found significant effects (the presentiment one actually found that higher quality studies fared better). Second, psi research has initiated or developed rigorous procedural and analytical strategies that mainstream psychology adopted later, and psi research is more rigorous in, for instance, using masked protocols, than psychology in general and other fields (Watt & Nagtegaal, 2004). Also, psi research has changed its procedures in response to internal and external criticisms, as exemplified by ganzfeld research. An analysis commissioned by the National Academy of Sciences concluded that alternative hypotheses (sensory leakage, recording or intentional errors, selective reporting, multiple analyses of variables, failures in randomization or statistical errors, and independence of studies) failed to explain away the significant effects in ganzfeld studies, which “regularly meet the basic requirements of sound experimental design” (Harris & Rosenthal, 1988, p. 53).

Nonetheless, some authors (e.g., Bösch et al., 2006; Rouder et al., 2013) have raised the possibility that supportive psi data could be due to nonpublication of failures to replicate. It is impossible to accurately know the potential effect of selective reporting, but psi research has taken steps for decades to reduce this possibility. For example, publication of nonreplications has been encouraged by journals for a long time (Broughton, 1987). In addition, this is such a small field that most researchers know who is researching what and can inquire about unpublished data to conduct meta-analyses. There are also known complete psi data sets that support the psi hypothesis (Baptista et al., 2015), and a psi critic wrote that selective publication is less evident in psi than in other areas (Hyman & Honorton, 1986). Furthermore, it should not be assumed that failures to replicate are not submitted for publication, whereas supportive experiments are. For instance, psi critics rushed to publish their failures to replicate Bem’s studies but not the supportive experiments in their database (dbem, 2012; Ritchie, Wiseman, & French, 2012). There have also been studies supportive of psi not submitted because they were conducted by skeptics (Sheldrake, 2015), or the researcher thought that there was already enough evidence for psi (Bem, personal communication, 2016). The selective publication effect cuts both ways and, when statistically evaluated in the reviewed meta-analyses, a file drawer effect has not been found to explain away the results, with the arguable exceptions of micro-PK and noncontact healing.

With regard to other questionable research practices (QRPs) such as “p hacking,” although one study showed that they were rampant in psychology (John, Loewenstein, & Prelec, 2012), another found that those results were probably inflated by the way the questions were phrased (Fiedler & Schwarz, 2016). A recent simulation of experimental psi data with a worst-case QRP scenario still found support for the psi hypothesis (Bierman, Spottiswoode, & Bijl, 2016), and psi researchers discussed QRP decades before the current debate (Office of Technology Assessment, 1989). Parapsychology has also taken steps to decrease potential QRP in the field through preregistration for psi research (Watt & Kennedy, 2015, 2017) and an open data registry (https://www.spr.ac.uk/publications/psi-open-data).

The argument about “exceptional claims requiring exceptional evidence,” although often adduced, is problematic for many reasons. They include the fact that many phenomena that we do not currently consider “exceptional” (e.g., electricity) were considered extraordinary if not impossible earlier in history and a requirement for “exceptional evidence” might have prevented them from becoming accepted. Then there is the problem of defining the criteria for “exceptional evidence,” which have varied across time, to a current proposed standard (Wagenmakers et al.’s [2011] prior Bayesian estimate for psi phenomena of 10\(^{-20}\)) that is virtually unfalsifiable. Deming (2016) concluded that there are not two different types of evidence in science and criticizes the misuse of the argument to “suppress innovation and maintain orthodoxy” (p. 1319). As Cornell professor in statistics (and psi skeptic) Joel Greenhouse (1991) stated, “parapsychologists should not be held to a different standard of evidence to support their findings than other scientists” (p. 388).

An additional point is that there is consistency across the meta-analyses and with descriptive research on psi phenomena. In both, awareness of putative psi phenomena often involves alterations in consciousness and salient emotional stimuli. The positive case for psi, however, should not be overstated because our knowledge of it is far from satisfactory and scientific conclusions are tentative. The level of replication, although comparable to other areas, leaves much to be desired, the ESs are small, and theories need to be developed and tested further.

Suggestions for Future Research

Psi research can be divided into studies that are mostly concerned with evaluating whether there is a psi effect to begin with (proof-oriented), and those that assume psi and seek to understand its nature (process-oriented; Stanford, 1974). Recommendations for process-oriented research are:
(a) Investigate the psychological and neurocognitive characteristics of people who perform well in controlled research. (b) Develop nonconscious measures of psi, following the promising results of presentment and implicit anomalous cognition research. (c) Conduct systematic research on naturally occurring instances of reputed psi in everyday life. (d) Develop and test strategies that might increase the ability to perform well in psi studies, for instance, through new procedures to elicit specific alterations of consciousness. (e) Analyze the characteristics of the research sessions/trials in which people were successful. (f) Investigate further the traits and practices of researchers who tend to be successful in their experiments.

With regard to proof-oriented research: (a) Use the project repositories already in existence and avoid QRP that place findings in doubt. This does not mean that serendipitous effects should be disregarded, but that they should be labeled as such. (b) Conduct studies that have enough power considering the ESs for that specific protocol. One important aspect, though, is to do it in such a way that the research does not become so tedious to experimenters and/or participants that the motivation goes down. (c) Develop and test procedures that approximate the characteristics of reputed psi in real life (e.g., Sheldrake, 2015). (d) Develop multidisciplinary protocols with researchers who are at least open to the reality of psi. Researchers who are already convinced that psi is impossible are likely to discourage participants (cf. Schlitz, Wiseman, Watt, & Radin, 2006). (e) Conduct prospective meta-analyses and standardize procedures following recent developments (Tressoldi & Utts, 2015, Utts was the 2016 president of the American Statistical Association). (f) Develop falsifiable theories to guide research. (g) Situate psi within larger domains such as the ongoing study of the general characteristics of consciousness/mind.

These proposals are unlikely to have much effect without a good infrastructure to support programmatic, well-funded research and theory, and serious endeavors should be valued rather than penalized. There should also be a change in the editorial policy of some journals so that the default position is not to automatically reject papers on psi but to have them evaluated on their own merits by knowledgeable and open reviewers. This would be in the spirit of William James and other founders of psychology, whose goal was to have a comprehensive and open discipline (Cardeña, Lynn, & Krippner, 2017).

The history of science shows that earlier certainties have been swept away by new findings and theories. This review of the empirical support for psi phenomena should make interested psychologists consult primary sources, their critiques (e.g., Wiseman, 2010), but also the responses to those critiques (e.g., Baptista & Derakhshani, 2014; Utts, 1991). After a careful investigation of the field, some readers might become convinced, as did the distinguished University of California neuroscientist James H. Fallon (2015) that at its best, psi research exhibits “methodological excellence” (p. xii) and promising results.

References


Received August 11, 2017
Revision received October 2, 2017
Accepted October 3, 2017