Journal of Parapsychology, 81(1), 9-32

INDIVIDUAL DIFFERENCE CORRELATES OF PSI PERFORMANCE IN FORCED-CHOICE PRECOGNITION EXPERIMENTS: A META-ANALYSIS (1945–2016)[1]

By Marco Zdrenka and Marc Stewart Wilson

ABSTRACT: Previous research in parapsychology has not been particularly persuasive, in large part due to a lack of replicability of significant findings. To address these concerns and better understand which factors may be associated with stronger and more consistent effect sizes, all forced-choice precognition experiments analysing individual differences (e.g., personality traits) were aggregated to determine which factors might reliably predict psi performance. Overall, 55 studies published between 1945 and 2016, including 35 individual difference measures, were subject to meta-analysis. Six individual difference measures, namely, luck belief (the belief that luck is primarily controllable), perceptual defensiveness, openness to experience, belief in psi, extraversion, and time belief as dynamic, were found to significantly correlate with psi performance. Given the particularly straightforward nature of forced-choice precognition experiments, a promising future avenue would be to explore these factors in confirmatory studies. It is hoped that researchers can model their future experiments off these findings in conjunction with preregistration techniques, to ultimately create a more systematic and robust database.

Keywords: meta-analysis, psi, precognition, personality, individual differences, forced-choice

Statistician Jessica Utts (1991) made the statement that a "promising direction for future process-oriented research [in parapsychology] is to examine the causes of individual differences in psychic functioning" (p. 377). It seems that such an approach is not only reasonable but also necessary, given that the evidence for psi (a general term used to describe anomalistic communication or interaction with the environment) is often inconsistent and elusive (Kennedy, 2001). If psi is to be taken seriously by the scientific community, its nature needs to be observable under prespecified conditions (Alcock, 2003; Hyman, 2010). Individual difference factors—such as specific personality traits (e.g., extraversion) or beliefs (e.g., belief in psi)—have been extensively analysed and thus represent a promising avenue in this regard. However, many researchers ignore individual difference factors (potentially missing important sources of between-individual variation in psi performance) or look at a multitude of varied factors that are difficult to sort through. An actual effect may also be masked if an individual difference factor is systematically related to psi performance. For example, participants who score high in a trait may overperform while participants who score low in that trait may underperform, effectively cancelling each other out. Therefore, this meta-analysis was intended to synthesise the relevant research to better understand the factors that may lead to a successful (or unsuccessful) and consistent demonstration of psi in the laboratory.

This meta-analysis focuses specifically on forced-choice experiments that have tested for precognition (i.e., the foreknowledge of an event without any known explanation). Forced-choice experiments give participants several options to choose from for their response, whereas free-choice experiments allow participants to make an unrestricted response. As free-choice experiments—such as the ganzfeld—have received a lot of attention in recent literature (see Bem & Honorton, 1994; Milton & Wiseman, 1999; Storm, 2006), this meta-analysis focuses exclusively on forced-choice experiments. It also focuses on precognition rather than telepathy (anomalous communication between people) or clairvoyance (perception without using normal sensory modalities), as precognition experiments are less susceptible to sensory leakage (Steinkamp, 2005). For example, in some telepathy experiments, participants may potentially make decisions based on the sender's or experimenter's facial cues, but this is not possible in precognition experiments in which the

[1] An earlier version of this paper was presented at the 10th Biennial European Conference of the Society for Scientific Exploration in Sigtuna, Sweden, 2016.

target cannot be known (even by the experimenter) until after the participant has already made their choice.

Previously, there have been two meta-analyses conducted on individual differences in psi laboratory research: one looking at extraversion (Honorton, Ferrari, & Bem, 1998) and the other looking at belief in psi, or what is known as the sheep-goat effect (Lawrence, 1993). Both meta-analyses found a relationship with psi performance (r = .09 and r = .03, respectively), indicating a small yet robustly significant overall effect size. However, many other individual differences have been examined in individual studies and it would be useful to summarise those studies here, and to compare them all with one another. Furthermore, many meta-analyses combine studies from multiple domains, making it difficult to unpack exactly what factors constitute a replicable psi experiment. Thus, the purpose of this meta-analysis is to (a) provide a comprehensive and updated review of all forced-choice precognition experiments that have included individual difference measures, and (b) estimate the overall magnitude of the relationship between each individual difference measure and psi performance, with the overall goal to provide researchers with the necessary information needed to design confirmatory studies.

Meta-Analysis

Although some meta-analyses have focused on only one research paradigm, such as the ganzfeld (Hyman, 1985) or biological systems (Schmidt, Schneider, Utts, & Walach, 2004), this is the first meta-analysis to combine both an experimental paradigm and individual differences. Forced-choice precognition is the chosen paradigm, as it is the most efficient method available for replication; the experiments are often automated (less potential for interference from both participants and experimenters); and exact probabilities of hits/misses can be objectively calculated. Although free-choice experiments can also be quantified, this requires an additional step, as participant responses need to be converted to target responses. This is avoided in forced-choice experiments altogether.

Although forced-choice precognition experiments might seem too narrow a subset to analyse, Steinkamp (2005) reviewed all forced-choice extrasensory perception (ESP) experiments—including telepathy, clairvoyance, and precognition—and found it difficult to come to any conclusions due to conflicting outcomes and wide variations in study designs. Furthermore, whereas some studies do not show any differences in effect sizes between precognition and other domains (see Steinkamp, Milton, & Morris, 1998), other studies have found a difference between clairvoyance, precognition, and telepathy effect sizes (Storm, Tressoldi, & Di Risio, 2012; Tart, 1983). Therefore, in defining the inclusion criteria narrowly, we sought to overcome the heterogeneity of studies in Steinkamp's (2005) review.

For the purposes of this meta-analysis, the effect size of interest is the correlation coefficient between the individual difference measure and psi performance—not psi performance specifically—with the participant as the unit of analysis.

Method

Retrieval of Studies

Only studies in the published literature are included in the meta-analysis, since parapsychology is a relatively small field and it is unlikely that there are many unpublished dissertations or theses (Honorton & Ferrari, 1989). Sourcing of relevant studies included the bibliography of two meta-analyses (Honorton & Ferrari, 1989; Storm et al., 2012), a database search (described below), along with an inspection of all English-language parapsychological journals, namely, the *Journal of Parapsychology*, *Journal of the American Society for Psychical Research*, *Journal of the Society for Psychical Research*, *Research in Parapsychology*, *Australian Journal of Parapsychology*, *European Journal of Parapsychology* (including the *Research Letter of the Utrecht University Parapsychology Laboratory*), and the *Journal of Scientific Exploration*.

An exhaustive search was conducted of research databases including PsycINFO, Google Scholar, WorldCat, and LexScien, using the keywords "individual differences," "precognition," "parapsychology," "forced-choice," "retrocausation," "retrocausality," "psi," "ESP," and "extrasensory perception." Most of

these searches located studies that were already found in the journals listed above.

The search period was intended to capture all experimental psi research published from 1945 through 2016 that included individual difference measures.

The search strategy revealed 35 individual difference variables including more common measures such as extraversion and belief in psi, along with less widely used measures such as temporal lobe dysfunction and latent inhibition.

Selection Criteria

Studies were included from 1945 until 2016 if they met the following criteria:

1. Forced-choice design
2. Precognition design
3. Included individual difference measure(s)
4. A minimum of two human participants

Studies that did not include relevant information were excluded. For example, the results reported in Wiseman and Greening (2002) could not be included as their precognition and clairvoyance data were combined when reporting individual difference measures (e.g., the sheep-goat effect), and results reported by Steinkamp (1998) could not be used as the number of participants was not reported.

The identification, screening, and eligibility of the studies followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al., 2009). Figure 1 provides a detailed summary of the database search and screening process.

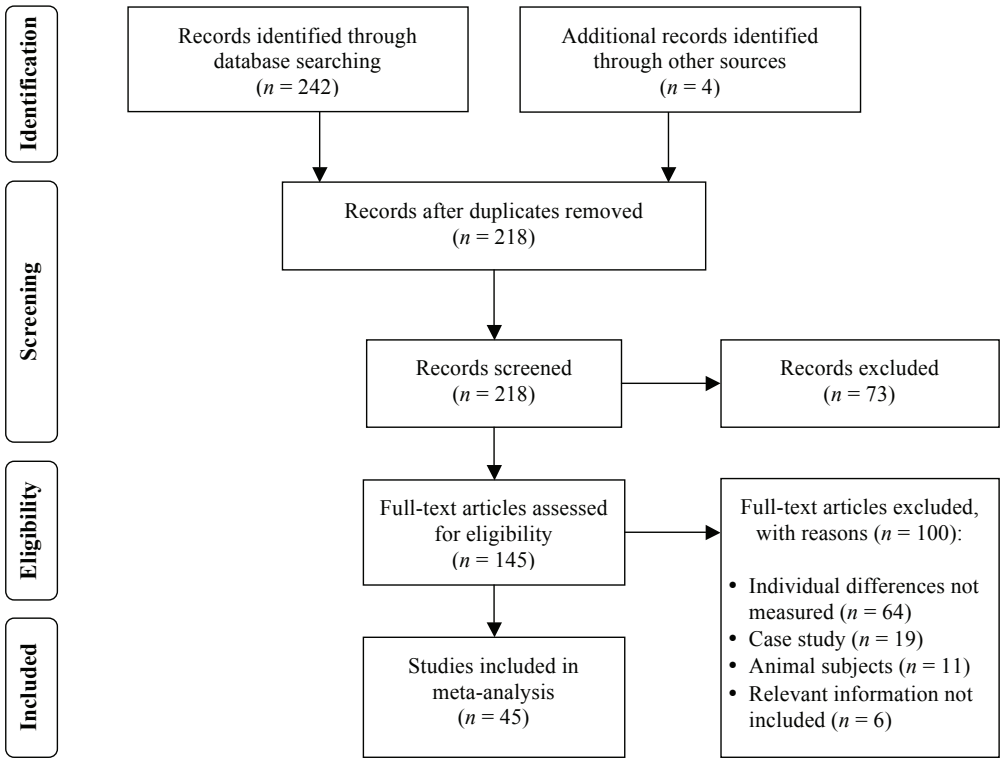

Figure 1. A flow diagram illustrating the database search and screening stages involved in this meta-analysis.[2]

[2] In both the screening stage and the eligibility stage, a number of articles (73 and 64, respectively) were excluded as they did not measure individual differences and were thus not relevant to this meta-analysis. The only difference is that the articles identified for exclusion in the screening stage were deemed irrelevant after only screening the titles, abstracts, and keywords, whereas the others required a more thorough reading of the article to make that determination.

Definitions

Independent investigator. For the purposes of this meta-analysis, an "investigator" refers to the lead author of a study. An investigator is considered independent of another investigator if both investigators have never worked on a paper together using the database being analysed (and/or worked with the other investigator's co-authors). The number of independent investigators can be helpful in determining how replicable an effect might be; the fewer independent investigators there are—even if the studies have been repeated multiple times—the less certain we can be that the results are replicable. As Hyman (1977) notes, it is not enough to simply repeat the same results, for whatever errors or biases may have occurred in the first instance might also be part of the subsequent repetitions of the experiment.

Individual differences. For the purposes of this research, an individual difference is defined as anything that an individual may psychologically vary on, whether it is personality, beliefs, intelligence, or aptitude (Nazimuddin, 2015). However, this meta-analysis makes a distinction between individual differences that are relatively constant regardless of situational factors (e.g., trait-level individual differences) and those that are more temporary and can be affected by the experimental situation (e.g., state-level individual differences such as participants' mood in the experiment); the latter (e.g., classification of a participant high in trait-anxiety but low in state-anxiety in a particular experimental setting) will not be included in this paper to reduce confusion while also limiting the number of variables analysed.

Individual differences were further categorised, combining similar measures (or subcomponents) into families of similar constructs. Where multiple measures of a single individual difference were used in a study, only the most appropriate measure was used. For example, some studies included both a sheep-goat measure and an interest in psi measure—in these cases, only the sheep-goat measure was included, as it is has historically been more consistently used as a measure of psi belief (Lawrence, 1993).

Procedural Features

As there are a multitude of individual difference measures included (35 individual difference constructs were identified for analysis), a separate meta-analysis was performed on each category of individual difference. The majority of these meta-analyses contain less than five studies in total, so it was not practical to code for procedural features.

Quality coding of the studies was not implemented for four reasons. Firstly, up until 1976, the founder of experimental parapsychology, J. B. Rhine, encouraged less detail in publications for nonsignificant parapsychological findings than significant findings (Steinkamp, 2005). Therefore, quality coding would inevitably favour newer studies, since more information is available for post-1976 studies (which may not correlate with the actual quality of the experiment). Secondly, precognition experiments have less potential for procedural defects compared with parapsychological research more generally, which is reflected by Honorton and Ferrari (1989) having only 6 quality criteria for precognition experiments rather than Rhine, Pratt, Stuart, Smith, and Greenwood's (1940) 34 quality criteria for ESP experiments. Thirdly, Honorton and Ferrari (1989) did not find a relationship between forced-choice precognition experiments and their quality ratings. Lastly, the majority of these meta-analyses had too few studies in total to meaningfully differentiate them on quality.

However, year of publication was coded, as it allowed us to examine whether effect sizes have increased over time, stayed the same, or even decreased. Honorton and Ferrari (1989) suggest that if effect sizes do not increase over time—as they found in their meta-analysis—it might mean that researchers lack an understanding of the underlying factors of psi performance (since they could not reliably increase its magnitude over time). Alternatively, an increase in effect size over time would be more promising, as was reported in Storm et al.'s (2012) meta-analysis.

Meta-Analysis of Correlation Coefficients

All indices of association between an individual difference measure and psi performance were converted to correlation coefficients using Comprehensive Meta-Analysis (CMA) software version 2 (Boren-

stein, Hedges, Higgins, & Rothstein, 2005) or manually. For example, *t* tests were converted to point-biserial correlations and phi coefficients were computed from 2 x 2 contingency tables. Some studies gave only trial-based data such as the critical ratio (*z*; e.g, Buzby 1967; Freeman & Nielsen, 1964). In these instances, correlations were estimated using a method for estimating effect sizes from critical ratios described by McCarthy and Schechter (1986), providing an estimate of Cohen's *d*—this was then converted to the *r* metric. Unreported correlations were estimated using the provided *p* values, whereas studies that reported only nonsignificance had their correlation set to .00,[3] a practice consistent with the approach adopted by Honorton et al. (1998) in their meta-analysis of extraversion and ESP performance. Where necessary, correlation signs were adjusted to reflect the appropriate relationship between the individual difference measure and psi performance. Finally, CMA weighted each study—using a random effects model incorporating both sample size and between-study variance—giving an overall outcome metric (*r*) for each individual difference measure in the database. A random effects model was used rather than a fixed effects model, as most studies were not exact replications of each other and this model takes into account such variation (Borenstein, Hedges, Higgins, & Rothstein, 2010). All *p* values are two-tailed.

Heterogeneity tests using Cochran's *Q* were also conducted on each meta-analysis to determine whether results from the included studies were representative of a single homogenous effect.[4] For those meta-analyses showing heterogeneity, moderator analyses were conducted using the year of publication as a proxy for methodological quality. The I^2 index was also reported, to give an idea of the degree of heterogeneity present. Finally, Rosenthal's (1979) fail safe *N*, or the file drawer estimate, was calculated for all meta-analyses that showed statistical significance to determine how many unreported studies averaging null results would need to exist for the effect to be reduced to overall nonsignificance. If the number is high, then there is less likelihood for publication bias, that is, studies being reported only if they show statistical significance (Honorton & Ferrari, 1989). Because unreported nonsignificant studies may have an average effect size below zero (Ferguson & Heene, 2012), an alternative method for examining publication bias, namely Egger's regression method, was also included for these studies (Egger, Smith, Schneider, & Minder, 1997). Egger's test aims to quantify potential asymmetrical distributions of studies around the mean effect size (Rothstein, 2008).

Results

Descriptives

Overall, this meta-analysis is comprised of 55 individual studies, which were reported in 45 papers and conducted by 17 independent investigators. The studies span a total of 71 years, between 1945 and 2016. There were a total of 17,584 participants analysed, with sample sizes ranging from 13 to 13,941. In the majority of these studies, students were the sample population.

Separate meta-analyses of the relationship between psi performance and each category of individual differences are reported below. It is ordered in terms of those variables that have the most exemplar studies (from belief in ESP, the Big Five, various operationalisations of luck, which have the most studies) through to variables for which there are only two or three studies (e.g., religiosity, emotional reactivity, intelligence).

For a summary of the total number of studies, independent investigators, and participants, see Table 2.

Major Individual Difference Measures

[3] However, this is an estimate, as the mean of the distribution of all possible nonsignificant outcomes is likely to be less than zero after removing the outcomes that give significant results by chance (as it effectively removes or truncates the right tail of the distribution). See the Appendix, Table A1, for all studies that this applies to.

[4] Note that these tests were conducted for all meta-analyses, even when there were only two or three studies, as it would be hard to justify an arbitrary cut-off point. However, it does not imply that all of these tests should be given equal weight. The heterogeneity analyses conducted on a limited number of studies should not be considered definitive.

Belief in ESP. Overall, belief in ESP was the most studied potential individual difference correlate in forced-choice precognition experiments, having been reported in 22 studies by 12 independent investigators based on a total of 2,200 participants. The most common measurement questionnaire was a variant of Schmeidler's (1943) sheep-goat criterion, such as Thalbourne and Delin's (1993) *Australian Sheep-Goat Scale* or Bahdra's (1966) *Sheep-Goat Questionnaire*. In general, participants who score high on these scales are classified as "sheep" (or believers in ESP) and those who score low are classified as "goats" (or disbelievers in ESP).

Figure 2 shows a forest plot of the correlation coefficients, with the correlations ranging from -.17 to .72. The overall mean weighted effect size (r) is .13 (p = .002), with a 95% confidence interval between .05 and .20. This suggests that there is a small but significant relationship between psi belief and performance on a psi task, such that people who believe in psi tend to perform better than those who do not believe in psi. This effect size is slightly larger than the effect size reported by Lawrence (1993) in his meta-analysis on the sheep-goat effect (r = .03), but that also included telepathy and clairvoyance experiments.

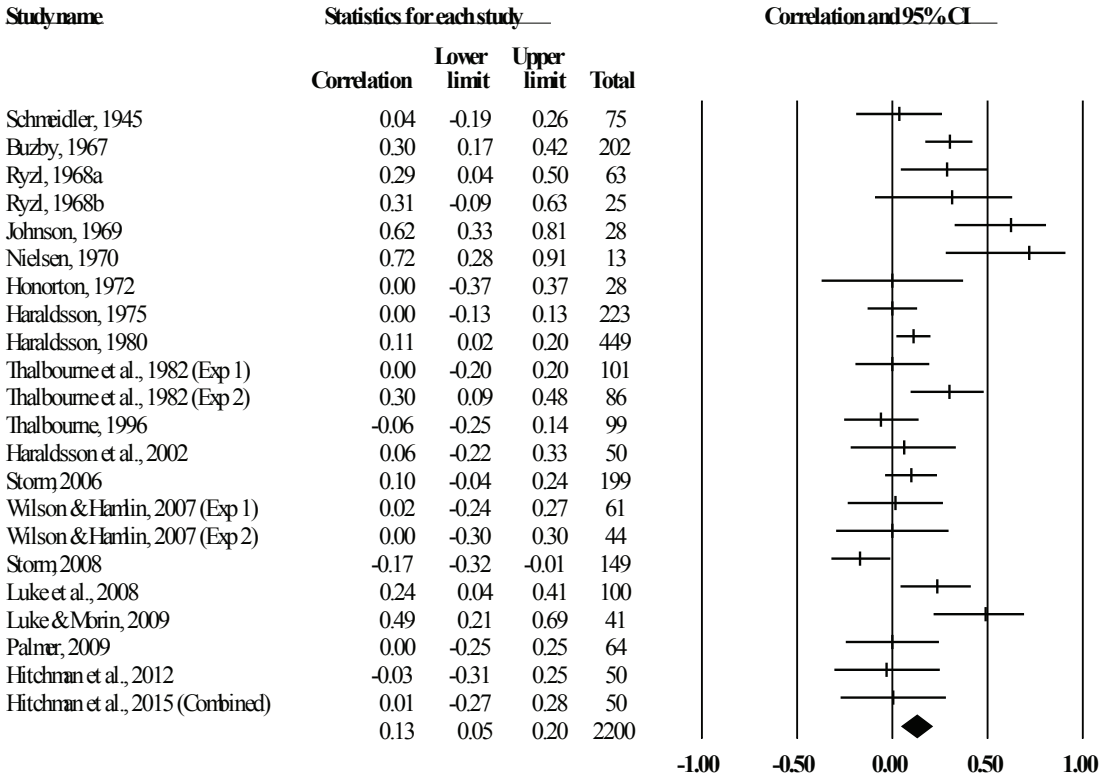

Figure 2. Meta-analysis of the relationship between belief in ESP and psi performance in forced-choice precognition experiments.

However, a test of heterogeneity was significant (Q = 61.16, p < .001) which suggests that variation in results may be due to factors other than the relationship between psi belief and performance (for example, error, or the influence of a moderator). The I^2 was 66%. Consequently, a mixed effects model (method of moments) meta-regression was conducted, which found year of publication to be a significant moderator, QR = 6.71, p = .01. Figure 3 shows effect sizes to decrease as year of publication increases. This means that some of the heterogeneity that caused the significant Q can be attributed to the year of publication.

Finally, the fail safe N, or the number of unreported studies averaging null results that would be needed to bring the p value to nonsignificance, is 141. Egger's test was not found to be significant, $t(20)$ = 1.21, p = .24.

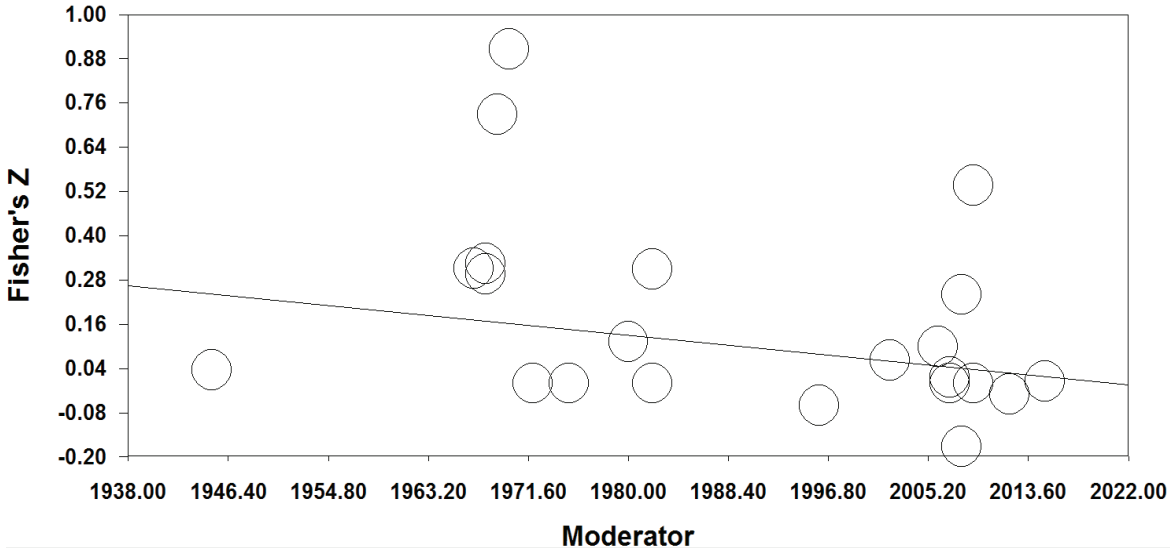

Figure 3. Meta-regression on the relationship between belief in ESP and psi performance in forced-choice precognition experiments, using publication date as the moderator.

The Big Five

The next set of meta-analyses cover personality indicators consistent with the Big Five personality traits (McCrae & Costa, 1987): extraversion (how outgoing and social a person is), neuroticism (a long-term tendency to be in a negative emotional state, such as anxious or frustrated), openness to experience (intellectually curious, willing to try new things, and imaginative), agreeableness (how kind and sympathetic a person is), and conscientiousness (organised and diligent).

Extraversion. Extraversion was the second most studied individual difference measure in forced-choice precognition experiments, having been reported in 14 studies by seven independent investigators with a total of 1,206 participants. Extraversion was typically measured within larger personality questionnaires such as the *16PF* (Cattell & Mead, 2008), but subcomponents such as Bem's (2011) *Sensation Seeking Scale* were also included. High scorers on these measures are generally considered to be extraverted and low scorers introverted.

Figure 4 shows a forest plot of the correlation coefficients, with the correlations ranging from -.28 to .35. The overall mean weighted effect size (r) is .08 (p = .02), with a 95% confidence interval between .01 and .15. This suggests that there is a small but significant relationship between extraversion and psi performance, such that people who are extraverted tend to perform better than those who are more introverted. This result is consistent with previous studies that have also found a positive relationship between extraversion and psi performance (Mangan, 1958; Palmer, 1978; Honorton et al., 1998). Furthermore, a test of heterogeneity was not significant (Q = 17.23, p = .19), with an I^2 index of 25%.

Finally, the fail safe N, or the number of unreported studies averaging null results that would be needed to bring the p value to nonsignificance, is nine, whereas Egger's test was not significant , $t(12)$ = 0.56, p = .59.

Neuroticism. Neuroticism was measured using a variety of different questionnaires encompassing anxiety, affect, and mood, and was included in nine studies by seven independent investigators and a total of 528 study participants. The correlation coefficients range from -.38 (Humphrey, 1945) to .60 (Freeman & Nielsen, 1964). The overall mean weighted effect size (r) is .05 (p = .43), with a 95% confidence interval between -.08 and .19. The results are inconclusive about whether an actual effect occurs, falling slightly short of Steinkamp's (2005) suggestion that neuroticism was a promising predictor of ESP forced-choice experiments.

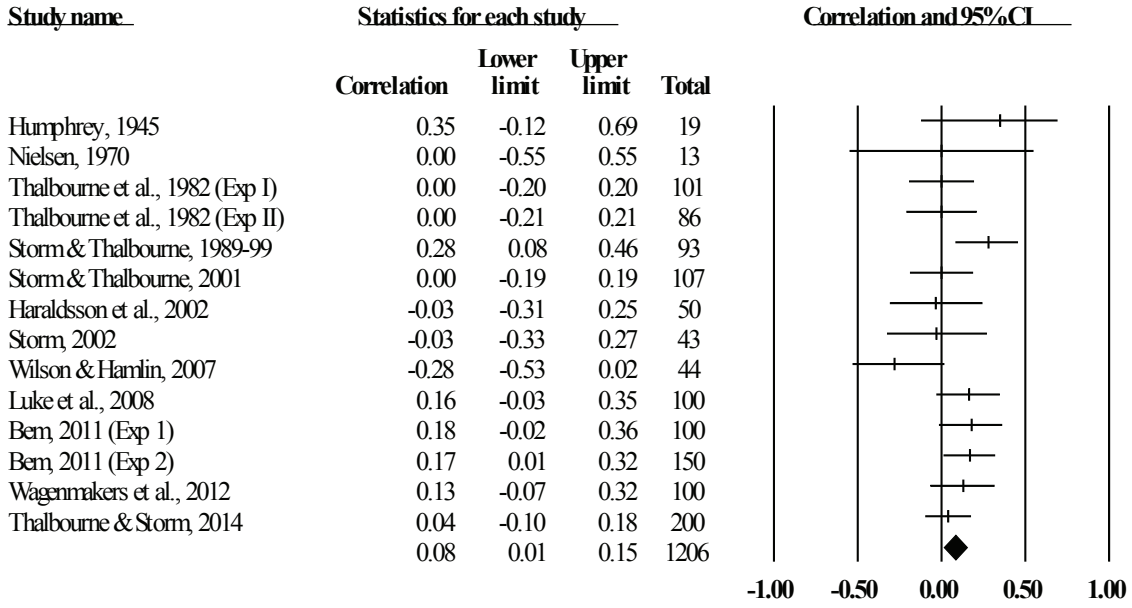

Figure 4. Meta-analysis of the relationship between extraversion and psi performance in forced-choice precognition experiments.

However, a test of heterogeneity was significant (Q = 17.29, p = .03) which suggests that there were potential moderating factors in this database. The I^2 was 54%. A mixed effects model (method of moments) meta-regression did not find year of publication to be a significant moderator, QR = 0.30, p = .58. Due to meta-regression analyses not being recommended for meta-analyses with less than 10 studies (Borenstein, Hedges, Higgins, & Rothstein, 2009), this finding should be treated with caution.

Openness to experience. Openness to experience was reported in nine studies of 522 participants, by five independent investigators. The most common measurement questionnaire was the *Openness to Experience Scale* (Goldberg, 1999). Figure 5 shows a forest plot of the correlation coefficients, with the correlations ranging from -.08 to .46. The overall mean weighted effect size (r) is .12 (p = .02), with a 95% confidence interval between .02 and .22, indicating a small but significant relationship between openness to experience and psi performance, such that people who prefer new experiences tend to perform better than those who prefer familiar routines. Furthermore, a test of heterogeneity was not significant (Q = 11.56, p = .24), with an I^2 index of 22%.

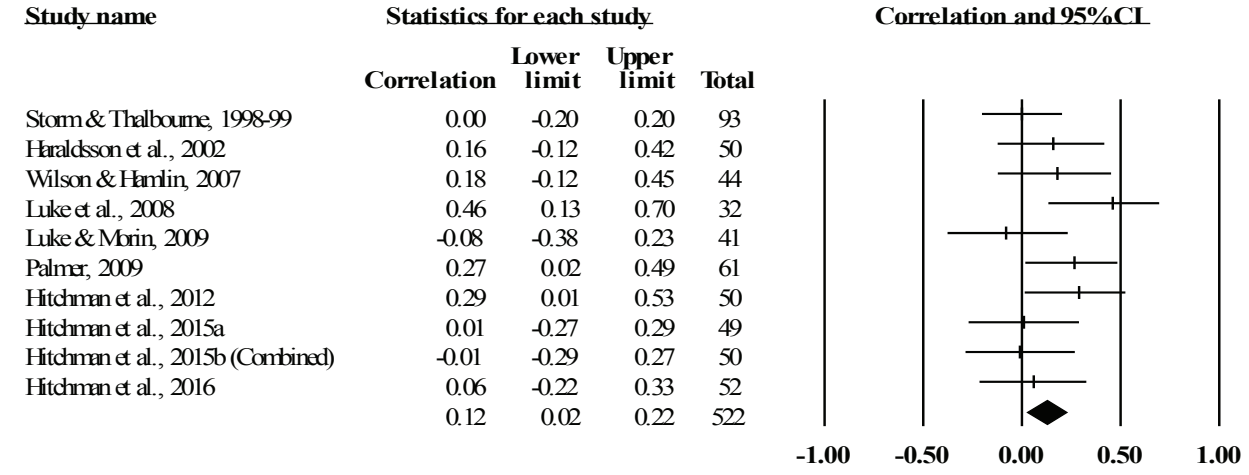

Figure 5. Meta-analysis of the relationship between openness to experience and psi performance in forced-choice precognition experiments.

In this case, the fail safe *N*, or the number of unreported studies averaging null results that would be needed to bring the *p* value to nonsignificance, is 12. Egger's test was not significant, $t(8) = 1.30$, $p = .23$.

Agreeableness. Agreeableness was reported in seven studies by four independent investigators with a total of 556 participants. The most common measurement questionnaire was the Independence factor of the *16PF* (Cattell, 1996; Cattell & Mead, 2008). Correlation coefficients range from -.36 (Humphrey, 1945) to .23 (Storm & Thalbourne, 1998-1999), with an overall mean weighted effect size (*r*) of .02 ($p = .71$) and a 95% confidence interval between -.09 and .13. Although the results are inconclusive, the data suggest that the true effect size is below .13 and could be zero. There was no indication of significant heterogeneity ($Q = 9.34$, $p = .16$), with an I^2 index of 36%.

Conscientiousness. Conscientiousness was the least studied Big Five personality trait in forced-choice precognition experiments, having been reported in only three studies by three independent investigators with a total of 187 participants. Conscientiousness was measured as a component of larger personality questionnaires such as the *NEO Five-Factor Inventory* (McCrae & Costa, 1987). Correlations range from .00 (Storm & Thalbourne, 1998-1999) to .23 (Wilson & Hamlin, 2007). The overall mean weighted effect size (*r*) is .06 ($p = .45$), with a 95% confidence interval between -.09 and .20, which is inconclusive but suggests that the effect size may lie below .20 and could be zero. Furthermore, a test of heterogeneity was not significant ($Q = 1.69$, $p = .43$). The I^2 is 0%.

Beliefs About Luck

Although luck has been explored in multiple studies, Smith (1998) showed that "luck" can mean different things to different people. Therefore, various measurements of luck and luck beliefs have been reported in forced-choice precognition experiments that include not only perceived luckiness, but also controllable luck belief, chance belief, providence belief, and fortune belief. The most common tool of measurement used in these experiments was Luke, Delanoy, and Sherwood's (2003) *Questionnaire of Beliefs about Luck*, which incorporates all of these subcomponents together. They will now be discussed in turn.

Perceived luckiness. Perceived luckiness has (prior to 2008) been the standard measurement used to explore luck in psi experiments and refers to how lucky one perceives oneself to be. For forced-choice precognition experiments, perceived luckiness was reported in four studies by two independent investigators with a total of 231 participants. Correlations range from -.20 (Hitchman, Row, & Sherwood, 2012) to .26 (Luke, Delanoy, & Sherwood, 2008), with an overall mean weighted effect size (*r*) of .08 ($p = .49$); 95% CI [-.14, .28]. These results are inconclusive but suggest that the effect size is below .28 (and could be zero). A test of heterogeneity was not significant ($Q = 7.09$, $p = .07$), with an I^2 index of 58%.

Luck belief. Luck belief refers to the belief that luck is primarily controllable, and participants who score high in this belief also view luck as internal, stable, and nonrandom (Luke et al., 2003). Luck belief was reported in five studies by one independent investigator with a total of 248 participants. Figure 6 shows a forest plot of the correlation coefficients, ranging from -.09 to .26. The overall mean weighted effect size (*r*) is .13 ($p = .048$), with a 95% confidence interval between .001 and .26, indicating a small but reliable relationship between luck belief and psi performance, such that people who believe luck to be controllable tend to perform better than those who see luck as uncontrollable. Furthermore, a test of heterogeneity was not significant ($Q = 4.11$, $p = .39$). The I^2 is 3%. Finally, the fail safe *N*, or the number of unreported studies averaging null results that would be needed to bring the *p* value to nonsignificance, is less than 1. However, Egger's test is not significant, $t(3) = 1.34$, $p = .27$.

Chance belief. Chance belief refers to the belief that luck is random, unpredictable, unstable, and inert (Luke et al., 2003). Chance belief was reported in five studies by one independent investigator with a total of 248 participants. Correlations range from -.16 (Luke et al., 2008) to .48 (Luke, Roe, & Davison, 2008). The overall mean weighted effect size (*r*) is .14 ($p = .23$), with a 95% confidence interval between -.09 and .36.

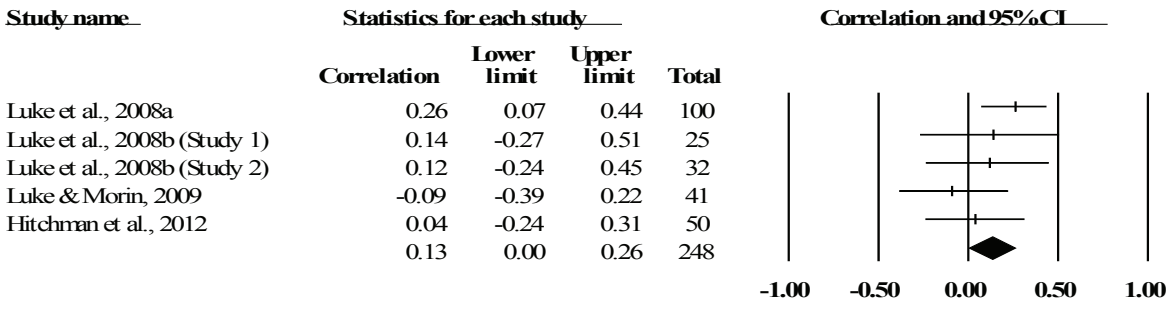

Figure 6. Meta-analysis of the relationship between luck belief and psi performance in forced-choice precognition experiments.

However, a test of heterogeneity was significant ($Q = 11.40$, $p = .02$) which suggests that there were potential moderating factors in this database. The I^2 is 65%. A mixed effects model (method of moments) meta-regression was conducted, but it did not find year of publication to be a significant moderator ($QR = 1.75$, $p = .19$).

Providence belief. Providence belief refers to the belief that luck is something that is managed by external forces or higher beings (Luke et al., 2003). Providence belief was reported in five studies by one independent investigator with a total of 248 participants. Correlations range from -.09 (Hitchman, Roe, & Sherwood, 2012) to .39 (Luke et al., 2008). The overall mean weighted effect size (r) is .12 ($p = .11$; 95% CI [-.03, .27]). Furthermore, a test of heterogeneity was not significant ($Q = 5.34$, $p = .25$). The I^2 is 25%.

Fortune belief. Fortune belief refers to the belief that luck is meant as a metaphor for life's successes rather than as a literal event (Luke et al., 2003) and was reported in five studies by one independent investigator based on 248 participants. Correlations range from -.08 (Luke & Morin, 2009) to .15 (Luke et al., 2008). The overall mean weighted effect size (r) is .03 ($p = .62$; 95% CI [-.10, .16]). A test of heterogeneity was not significant ($Q = 1.98$, $p = .74$), and the I^2 index is 0%.

Uncategorised Individual Difference Measures

Creativity. Creativity was reported in nine studies by three independent investigators with a total of 506 participants. The most common measurement questionnaires were the short version of the *Personal-Social Motivational Inventory* (Torrance, 1963) and the *Creative Cognition Inventory* (Holt, 2002). Correlations range from -.17 (Schmeidler, 1964c) to .20 (Luke et al., 2008), with an overall mean weighted effect size (r) of .05 ($p = .46$; 95% CI [-.08, .17]). Furthermore, a test of heterogeneity was not significant ($Q = 5.90$, $p = .21$), with an I^2 index of 32%.

Perceptual defensiveness. Perceptual defensiveness refers to psychological defence mechanisms and is related to subliminal perception and preconscious processing. All reported studies administered the *Defense Mechanism Test* (Kragh & Smith, 1970), with a total of six studies conducted by one independent investigator and a total of 272 participants. The test incorporates a tachistoscopic technique using peripheral stimuli to trigger subliminal anxiety and thereby defensive reactions. Figure 7 shows a forest plot of the correlation coefficients, with the correlations ranging from -.04 to .30. The overall mean weighted effect size (r) is .12 ($p = .049$; 95% CI [.001, .24]), suggesting a small but significant relationship between perceptual defensiveness and psi performance, such that people who exhibit high preconscious defensiveness tend to perform better than those who do not. There was no evidence of heterogeneity ($Q = 5.13$, $p = .40$), with an I^2 index of 3%.

Transliminality. Transliminality is defined as "the hypothesised tendency for psychological material to cross thresholds into or out of consciousness" (Thalbourne & Delin, 1994, p. 31), and was used in five studies by one independent investigator based on a total of 542 participants, with the most common measurement questionnaire being the *Transliminality Scale* (Thalbourne, 1998). Correlations range from

-.13 (Thalbourne, 1996) to .27 (Storm & Thalbourne, 1998-1999). The overall mean weighted effect size (*r*) is .01 ($p = .91$), with a 95% confidence interval between -.13 and .15. Although these results are inconclusive, a test of heterogeneity was significant ($Q = 9.81$, $p = .04$), but year of publication was not a significant moderator ($QR = 1.29$, $p = .26$). Due to the small number of studies, this finding should be treated with caution (Borenstein et. al., 2009). The I^2 is 59%.

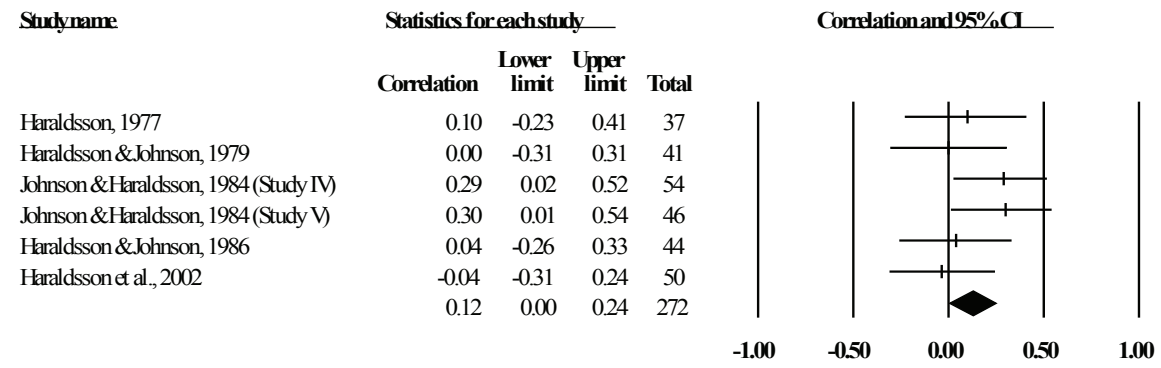

Figure 7. Meta-analysis of the relationship between perceptual defensiveness and psi performance in forced-choice precognition experiments.

Pro attitude. Pro attitude refers to "an attitude that is favourably directed towards an outcome" (Storm, 2002, p. 47) and three studies with 393 participants were conducted by a single investigator. All used Thalbourne and Storm's (2014) *Pro Attitude Scale*. Correlations range from -.16 (Storm, 2002) to .02 (Storm, 2008), with an overall mean weighted effect size (*r*) of -.02 ($p = .70$; CI 95% [-.12, .08]). A test of heterogeneity was not significant ($Q = 1.04$, $p = .60$), with an I^2 of 0%.

Dream recall. Dream recall (specifically, whether an individual recalls their dreams or not) was reported in four studies by three independent investigators with a total of 799 participants. It was typically measured using a one-item questionnaire, which asked participants how frequently they recalled their own dreams. Correlations range from .03 (Thalbourne, 1996) to .43 (Honorton, 1972). The overall mean weighted effect size (*r*) is .07 ($p = .23$; 95% CI [-.04, .18]), and a test of heterogeneity was not significant ($Q = 5.84$, $p = .12$). The I^2 is 49%.

Reports of unusual spontaneous experiences. Unusual spontaneous experiences can be described as seemingly paranormal experiences in everyday life (as opposed to the experimental laboratory) and were reported in three studies by two independent investigators based on a total of 695 participants. It was typically measured using a single question, which asked participants if they have had any precognitive dreams (i.e., dreams that they thought predicted the future). Correlations range from -.21 (Schmeidler, 1964d) to .00 (Haraldsson 1975, 1980), with an overall mean weighted effect size (*r*) of -.01 ($p = .87$; 95% CI [-.08, .07]). A test of heterogeneity was not significant ($Q = 0.85$, $p = .65$), with an I^2 of 0%.

Religiosity. Religiosity was reported in two studies by two independent investigators with a total of 149 participants. In both studies, religiosity was measured using the *Religiosity Scale* (Haraldsson, 1993), with the correlations ranging from -.13 (Thalbourne, 1996) to .08 (Haraldsson, Houtkooper, Schneider, & Bäckström, 2002). The overall mean weighted effect size (*r*) is -.05 ($p = .59$; 95% CI [-.24 and .14]). A test of heterogeneity was not significant ($Q = 1.34$, $p = .25$). The I^2 is 26%.

Emotional reactivity. Emotional reactivity is a measure of one's emotional reaction to violent, scary, or gruesome content in photographs, movies, and videos. All studies used the *Emotional Reactivity Scale* (Bem, 2003). A total of three studies by one investigator looked at emotional reactivity, with 151 participants being included in the experiments. Correlations range from -.27 (Hitchman, Sherwood, & Roe, 2015) to .29 (Hitchman, Pfeuffer, Roe, & Sherwood, 2016), with an overall mean weighted effect size (*r*) of .06 ($p = .71$; 95% CI [-.27, .38]). A test of heterogeneity was significant ($Q = 8.53$, $p = .01$), with the year of publication being a significant moderator ($QR = 3.96$, $p = .046$). Figure 8 shows effect sizes to increase as year of publication increases. The I^2 is 77%.

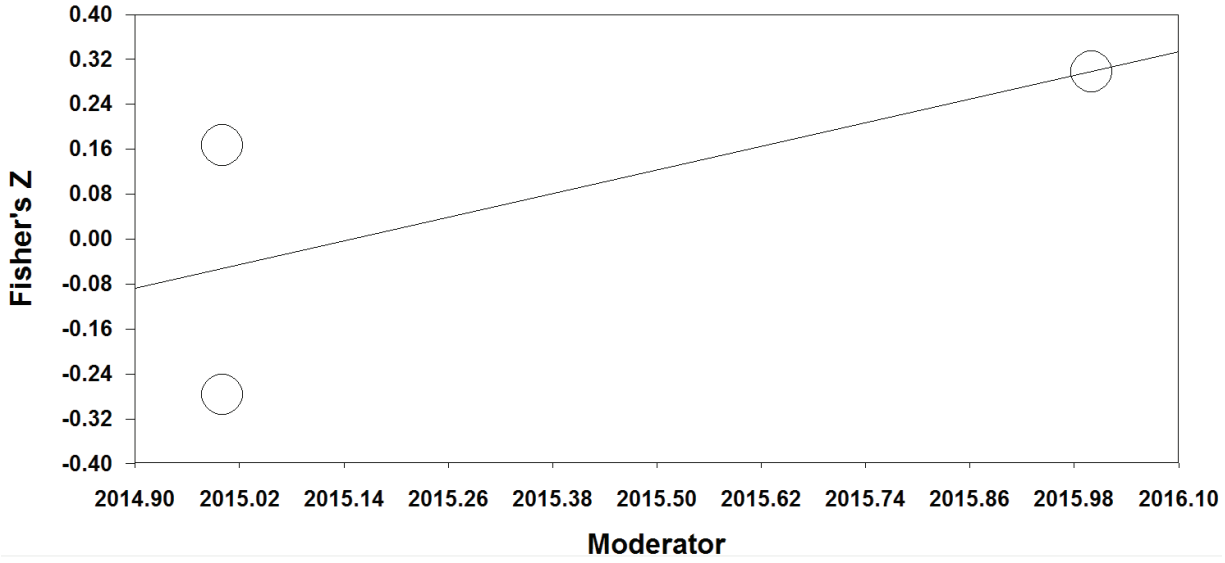

Figure 8. Meta-regression on the relationship between emotional reactivity and psi performance in forced-choice precognition experiments, using publication date as the moderator.

Temporal lobe dysfunction. Temporal lobe dysfunction measures symptoms of temporal lobe damage such as disturbances of perception, selective attention of auditory input, and impaired organisation of verbal material. A total of two studies by two independent investigators looked at temporal lobe dysfunction, using either the 13-item *LIMBEX Scale* or the Complex Partial Epileptic Signs cluster of the *Personal Philosophy Inventory* (Persinger & Makarec, 1987). There were a total of 114 participants in all of the experiments. Correlations range from -.01 (Hitchman, Roe, & Sherwood, 2015) to .00 (Palmer, 2009), with an overall mean weighted effect size (r) of -.004 (p = .96; 95% CI [-.19, .18]). A test of heterogeneity was not significant (Q = 0.003, p = .96), with an I^2 of 0%.

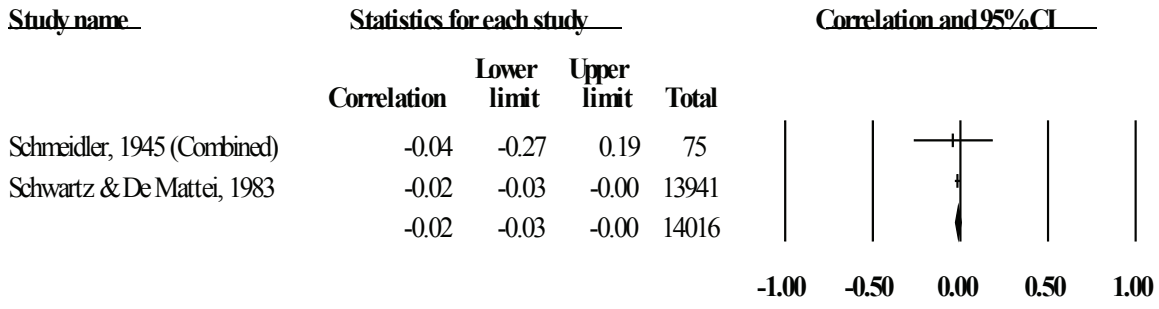

Figure 9. Meta-analysis of the relationship between time belief as dynamic and psi performance in forced-choice precognition experiments.

Time belief as dynamic. "Time belief as dynamic" was reported in two studies by two independent investigators with a total of 14,016 participants and refers to how strongly one sees time as being in a constant flow. In both studies, Knapp and Garbutt's (1958) *Time Metaphor Test* was administered, which measures the belief that time is metaphorically "a dashing waterfall" or "a fast moving shuttle," for example. Figure 9 shows a forest plot of the correlation coefficients, with the correlations ranging from -.04 to -.02. The overall mean weighted effect size (r) is -.02 (p = .04; (95% CI [-.03, -.001]), suggesting a small but significant negative relationship between time belief as dynamic and psi performance, such that people who do not view time as dynamic tend to perform better than those who see time as dynamic and hasty.

Furthermore, a test of heterogeneity was not significant ($Q = 0.05$, $p = .82$), and the I^2 is 0%. The fail safe N was not calculated, as it is impractical to do so with less than three studies.

Future-orientation. Future-orientation refers to being more attentive towards future events than past events and was reported in two studies by two independent investigators based on a total of 118 participants. Both studies used the *Attitude Toward the Future Questionnaire* (Vaughan & Houck, 1993). Correlations range from -.04 (Haraldsson et al., 2002) to .12 (Vaughan & Houck, 1993), and the overall mean weighted effect size (r) is .05 ($p = .57$; 95% CI [-.13, .23]). A test of heterogeneity was not significant ($Q = .65$, $p = .42$), with an I^2 of 0%.

Intelligence. Intelligence was measured in two studies by two independent investigators with a total of 80 participants. Both studies included only child participants, ranging in age from 4 to 14. The measurement tools used were the *Peabody Picture Vocabulary Test* (Dunn & Hottel, 1961) and a formal mathematical ability test using basic addition, subtraction, multiplication, and division problems. Correlations range from -.23 (Winkelman, 1981) to .04 (Drucker, Drewes, & Rubin, 1977), with an overall mean weighted effect size (r) of -.07 ($p = .62$; 95% CI [-.31, .19]). A test of heterogeneity was not significant ($Q = 1.2$, $p = .27$), and the I^2 is 19%.

Single Studies

There were also a number of individual difference measures that were analysed only for a single study. As single studies cannot be meta-analysed, they will be presented individually in Table 1 below, in order of effect size.

Table 1

Summary Statistics of All Individual Difference Measures Included in Only One Study

Individual Difference Measure	Study Author(s)	Effect Size (r)	Sample Size
Memory	Stanford, 1970	.36*	30
Sensitivity to Punishment	Hitchman et al., 2016	-.21	51
Optimism	Haraldsson et al., 2002	.20*	50
Fantasy Proneness	Thalbourne, 1996	-.13	99
Latent Inhibition	Hitchman , Sherwood, & Roe, 2015	-.11	48
Belief in the Occult	Haraldsson et al., 2002	.10	50
Time Belief as Naturalistic	Schmeidler, 1964b	-.08	75
Psychotism	Haraldsson et al., 2002	-.08	50
Hypersensitivity	Thalbourne, 1996	.04	99
Time Belief as Humanistic	Schmeidler, 1964b	.01	75
Cerebral Lateralization	Palmer, 2009	.00	64
Sensitivity to Reward	Hitchman et al., 2016	.00	51

*$p < .05$

Summary of Results

Below (Table 2) is a summary of all of the individual difference measures that were meta-analysed for comparison. The individual difference measures are ordered by overall effect size.

Column 1 is the individual difference measure; Column 2 is the 95% confidence interval,[5] Column 3 is the overall effect size; Column 4 is the number of individual studies included in the meta-analysis; Column 5 is the number of independent investigators in the meta-analysis; Column 6 is total number of participants for all of the studies in the meta-analysis combined.

Table 2
Summary Statistics of All Individual Difference Measures That Were Meta-Analysed

Individual Difference Measure	95% Confidence Intervals	Overall Effect Size (*r*)	Number of Studies	Number of Independent Investigators	Total Number of Participants
Chance Belief	-.09, .36	.141	5	1	248
Luck Belief	.001, .26	.131*	5	1	248
Belief in Psi	.05, .20	.125*	22	12	2,200
Perceptual Defensiveness	.001, .24	.125*	6	1	272
Providence Belief	-.03, .27	.125	5	1	248
Openness to Experience	.02, .22	.124*	9	5	522
Extraversion	.01, .15	.080*	14	7	1,206
Perceived Luckiness	-.14, .28	.076	4	2	231
Dream Recall	-.04, .18	.070	4	3	799
Intelligence	-.31, .19	-.065	2	2	80
Emotional Reactivity	-.27, .38	.064	3	1	151
Conscientiousness	-.09, .20	.056	3	3	187
Neuroticism	-.08, .19	.054	9	7	528
Religiosity	-.24, .14	-.054	2	2	149
Future-Oriented	-.13, .23	.053	2	2	118
Creativity	-.08, .17	.047	9	3	506
Fortune Belief	-.10, .16	.032	5	1	248
Agreeableness	-.09, .13	.021	7	4	556
Pro Attitude	-.12, .08	-.019	3	1	393
Time Belief as Dynamic	-.03, -.001	-.017*	2	2	14,016
Transliminality	-.13, .15	.008	5	1	542
Spontaneous Experences	-.08, .07	-.006	3	2	695
Temporal Lobe Dysfunction	-.19, .18	-.004	2	2	114

*$p < .05$

Discussion

As we can see from the summary of results, the majority of these individual difference measures have not been extensively investigated, with the exception of belief in psi, extraversion, and neuroticism. It

[5] As effect size estimates based on previous research are inherently uncertain, confirmatory studies based on lower confidence intervals are less likely to overestimate the true effect size (Kennedy, 2016). Therefore 80% and 68% confidence intervals are provided in the Appendix, Table A2.

might therefore be argued that such a meta-analysis is unnecessary. However, without a meta-analysis, researchers will likely impose their own synthesis of the data, and a meta-analysis can provide greater clarity in this regard—even if it only incorporates two or three studies, as Valentine, Pigott, and Rothstein (2010) argue that all other synthesis techniques are less transparent and/or less likely to be valid. At the same time, it is not intended to stop researchers from exploring individual difference measures that may not yet seem promising, especially those that have been tested only a handful of times; it is merely given as a benchmark of past results.

With that being said, the results suggest that there may be only a small pool of individual difference measures that are robustly correlated with performance on a forced-choice precognition task. This is also consistent with Steinkamp's (2005) review of forced-choice ESP experiments, where she found that "there are few variables which have correlated clearly with success . . . most variables tested provided little evidence either way as being ultimately psi-conducive and there were relatively few variables that appeared to be encouraging" (p. 155). However, notable exceptions in this meta-analysis include extraversion and belief in psi, which show more consistent results across a larger number of studies. It should also be noted that with the number of meta-analyses conducted in this paper, there is an increased risk of family-wise error, and that one or more of these significant findings might be the result of multiple analyses (e.g., represent a false positive). Further, while forced-choice ESP tests produce normal distributions (unless the number of trials is very small and/or the number of response alternatives is very large), it may be useful to incorporate nonparametric statistics in the case of any highly skewed distributions. Researchers should bear this in mind when setting their expectations for future experiments.

Nevertheless, the findings suggest a small but significant relationship between the following individual difference measures and psi performance: luck belief (specifically, the belief that luck is primarily controllable), perceptual defensiveness, openness to experience, belief in psi, extraversion, and time belief as dynamic. Perhaps what underlies these individual differences is a mechanism derived from being open-minded, curious, social, and intuitive—all of which might lead people to discuss, think about, and explore the "paranormal." Consequently, these same people may act on information or intuitions that others may ignore or supress, leading them to make better decisions about the future than we would expect by chance alone. This may mean that a relationship exists between these variables and performance on a psi task, where such traits may either facilitate or innately allow demonstrations of psi.

However, given that even the strongest predictor (extraversion) in this meta-analysis accounts only for approximately 2% of the variance on its own, perhaps these predictors are not related and are instead additive, and provide more power when analysed together. Therefore, the best strategy for future researchers may be to combine individual difference factors, not just for the additive benefits but also to examine potential interactions (see Baron & Kenny, 1986) between the factors that may predict precognitive performance.

Alternatively, these results may be due to statistical anomalies, having arisen from the large amount of analyses being conducted on individual differences in psi research, or even due to methodological flaws. The other possibility is that the results reflect an actual relationship between certain variables combined with imperfect research designs. Taking into consideration the fact that many of these findings, including the nonsignificant results, were based on only a limited amount of studies, it is difficult to come to any strong conclusions. If one also considers the possibility of experimenter psi (i.e., where the experimenter influences the final results of an experiment due to his or her own psi abilities), it becomes extremely difficult to disentangle the data, especially in the case of a meta-analysis with only a single independent investigator. This potential explanation has previously been offered for the relationship between perceptual defensiveness and psi performance (see Haraldsson et al., 2002).

Retrospective meta-analyses also have several limitations, so it is not appropriate to make any definitive statements about the results without first conducting confirmatory studies. One such limitation of retrospective meta-analysis is that included studies are often affected by publication bias or the file-drawer effect, whereby only significant results are reported or published. Although parapsychological journals generally publish more nonsignificant results than most mainstream scientific journals (Mousseau, 2003), no field is entirely immune, especially when there may be tens or even hundreds of secondary analyses con-

ducted (e.g., various individual difference measures). Indeed, the low fail safe *N* numbers found in several of these meta-analyses (e.g., nine for extraversion) suggest that publication bias is a possibility. At the same time, there is no indication to argue strongly that publication bias is a problem when taking into account Egger's test results, which should be reassuring for parapsychologists given Mousseau's (2003) findings. Secondly, there will always be subjectivity involved in meta-analytical procedures and interpretation, such as defining and judging exclusion criteria, using search strategies, or coding the studies (Murray, 2011). Biases will come into play—whether conscious or unconscious—that influence procedural decisions, especially since researchers will already be aware of the results of the individual studies (Watt & Kennedy, 2016). This subjectivity allows psi proponents and critics to conduct meta-analyses whose conclusions often support their own prior beliefs, but never manage to convince the other side (Palmer, 2003).

Yet meta-analyses are still useful in that they can suggest the conditions under which replication is most likely to occur, assuming an effect exists at all. An overall effect size also gives future researchers the ability to calculate how many participants they would need to include in their experiment for it to be adequately powered. With these key pieces of information, prospective meta-analyses (which define the exclusion criteria and other details beforehand) can then be conducted using only future studies that are to be preregistered. A prospective meta-analysis therefore avoids all the potential issues of publication bias and subjectivity that are evident in a retrospective meta-analysis, while also addressing methodological issues such as optional stopping. In fact, Watt (2016) has recently set up a registration-based prospective meta-analysis of one of the most thorough yet controversial paradigms in parapsychology, the ganzfeld—Watt's (2016) meta-analysis protocol specifically includes only preregistered individual studies that prospectively fit their criteria. This is a positive direction for parapsychology, as it brings structure and focus to the field.

Preregistration has been made even easier by the Koestler Parapsychology Unit (KPU) registry, an initiative started at the University of Edinburgh in 2012 that allows researchers to prospectively register their experiments in detail, publically, and is not affiliated with a specific journal (Watt & Kennedy, 2015). Not only are prospective meta-analyses the ideal way to test the replicability of psi phenomena, but they are also the best way to confirm the null hypothesis should psi not exist. Alcock (2003) claims that the latter hypothesis often does not get serious consideration by parapsychologists, so multiple prospective meta-analyses showing nonsignificance may force parapsychologists to give the null hypothesis more deliberation than a single study or retrospective meta-analysis would.

Another consideration for attempting replication is the researcher conducting the experiment. Although some parapsychologists believe that the psi experimenter effect eliminates the possibility of true replication, that is, that due to the nature of psi only experimenters who are proponents of psi will get positive results in psi experiments whereas sceptics will not (Utts, 2015), most researchers would only be satisfied that psi phenomena exist if it were to be consistently demonstrated by neutral scientists and not just a select few who believe in psi (Alcock, 2003; Palmer, 2016). The current meta-analysis was conducted with this goal in mind, as forced-choice precognition experiments are arguably the easiest to run and can be automated using computer programmes. For example, Bem (2011) ran his Precognitive Detection of Erotic Stimuli experiment using an automated computer programme. This allows researchers to collect large amounts of data with relatively little effort, an important consideration if researchers are to try and replicate the small effect sizes shown in this meta-analysis (Steinkamp, 2005).

Ultimately, it is hoped that this meta-analysis can be used as a springboard for future research, allowing the findings to be used in a productive way and perhaps aiding in the development of research programmes that are specific and structured. As Watt (2005) comments, "Parapsychologists need to be far more systematic in how they tackle these questions. . . . Systematic follow-up is an essential prerequisite for demonstrating a replicable effect" (p. 222). With parapsychology being such a small field, it is important that researchers work together to build up a body of evidence that is considered respectable by both parapsychologists and mainstream academics. With the recent failures to replicate many foundational studies in both psychology and medicine (Open Science Collaboration, 2015), now is the perfect time to define what a replicable psi experiment really is and take advantage of the benefits of preregistration. Only then will

we be able to finally confirm or disconfirm some of the major hypotheses in psi research. Depending on whether you are extraverted or believe in psi, you may already know how it will turn out.

References

[*Paper in the meta-analysis]

Alcock, J. (2003). Give the null hypothesis a chance: Reasons to remain doubtful about the existence of psi. *Journal of Consciousness Studies*, *10*, 29–50.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. doi:10.1037/0022-3514.51.6.1173

Bem, D. J. (2003). Precognitive habituation: Replicable evidence for a process of anomalous cognition. *Proceedings of Presented Papers: The Parapsychological Association 46th Annual Convention*, 6–20.

*Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407–425. doi:10.1037/a0021524

Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, *115*, 4–18. doi:10.1037/0033-2909.115.1.4

Bhadra, B. H. (1966). The relationship of test scores in ESP. *Journal of Parapsychology*, *30*, 1–17.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive meta-analysis (version 2) [Computer software]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to Meta-Analysis*. Chichester, UK: Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*, 97–111. doi:10.1002/jrsm.12

*Buzby, D. E. (1967). Subject attitude and score variance in ESP tests. *Journal of Parapsychology*, *31*, 43–50.*

Cattell, H. E. P. (1996). The original big-five: A historical perspective. *European Review of Psychology, 46*, 5–14.

Cattell, H. E., & Mead, A. D. (2008). The Sixteen Personality Factor Questionnaire (16PF). In G. Boyle, G. Matthews, & D. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment* (Vol. 2), (pp.135–178). Trowbridge, UK: Cromwell Press.

*Drucker, S. A., Drewes, A. A., & Rubin, L. (1977). ESP in relation to cognitive development and IQ in young children. *Journal of the American Society for Psychical Research, 71*, 289–298.

Dunn, L. M., & Hottel, J. V. (1961). Peabody Picture Vocabulary Test performance of trainable mentally retarded children. *American Journal of Mental Deficiency, 65*, 448–452.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 629–634. doi: 10.1136/bmj.315.7109.629

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, *7*, 555–561. doi:10.1177/1745691612459059

*Freeman, J., & Nielsen, W. (1964). Precognition score deviations as related to anxiety levels. *Journal of Parapsychology, 28*, 239–249.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe, 7*, 7–28.

*Haraldsson, E. (1975). Reported dream recall, precognitive dreams, and ESP [Abstract]. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1974* (pp. 47–48). Metuchen, NJ: Scarecrow Press.

*Haraldsson, E. (1980). Scoring in a precognition test as a function of the frequency of reading on psychical phenomena and belief in ESP. *Research Letter, 1,* 1–8.

Haraldsson, E. (1993). Are religiosity and belief in an afterlife better predictors of ESP performance than belief in psychic phenomena?. *Journal of Parapsychology*, *57*, 259–273.

*Haraldsson, E., Houtkooper, J. M., Schneider, R., & Bäckström, M. (2002). Perceptual defensiveness and ESP performance: Reconstructed DMT ratings and psychological correlates in the first German DMT-ESP experiment. *Journal of Parapsychology, 66,* 249–270.

*Hitchman, G. A., Pfeuffer, C. U., Roe, C. A., & Sherwood, S. J. (2016). The effects of experimenter-participant interaction qualities in a goal-oriented nonintentional precognition task. *Journal of Parapsychology*, *80*, 45–69.

*Hitchman, G. A., Roe, C. A., & Sherwood, S. J. (2012). A reexamination of nonintentional precognition with openness to experience, creativity, psi beliefs, and luck beliefs as predictors of success. *Journal of Parapsychology*, *76*, 109–145.

*Hitchman, G. A., Roe, C. A., & Sherwood, S. J. (2015). The relationship between lability and performance at inten-

tional and nonintentional versions of an implicit PMIR-type psi task. *Journal of Parapsychology, 79*, 65–86.
*Hitchman, G. A., Sherwood, S. J., & Roe, C. A. (2015). The relationship between latent inhibition and performance at a non-intentional precognition task. *Explore: The Journal of Science and Healing*, *11*, 118–126. doi:10.1016/j.explore.2014.12.004
Holt, N. J. (2002). *Creativity and boundary permeability*. Unpublished manuscript, University of Northampton, UK.
*Honorton, C. (1972). Reported frequency of dream recall and ESP. *Journal of the American Society for Psychical Research*, *66*, 369–374.
Honorton, C., & Ferrari, D. C. (1989). "Future telling": A meta-analysis of forced-choice precognition experiments, 1935-1987. *Journal of Parapsychology*, *53*, 281–308.
Honorton, C., Ferrari, D. C., & Bem, D. J. (1998). Extraversion and ESP performance: A meta-analysis and a new confirmation. *Journal of Parapsychology*, *62*, 255–276.
*Humphrey, B. M. (1945). An exploratory correlation study of personality measures and ESP scores. *Journal of Parapsychology*, *9*, 116–123.
Hyman, R. (1977). The case against parapsychology. *The Humanist, 37*, 37–49.
Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology, 49*, 3–49.
Hyman, R. (2010). Meta-analysis that conceals more than it reveals: Comment on Storm et al. (2010). *Psychological Bulletin, 136*, 486–490. doi:10.1037/a0019676
Kennedy, J. E. (2001). Why is psi so elusive? A review and proposed model. *Journal of Parapsychology, 65*, 219–246.
Kennedy, J. E. (2016). Is the methodological revolution in psychology over or just beginning? *Journal of Parapsychology*, 80, 156–168.
Knapp, R. H., & Garbutt, J. T. (1958). Time imagery and the achievement motive. *Journal of Personality*, *26*, 426–434. doi:10.1111/j.1467-6494.1958.tb01597
Kragh, U., & Smith, G. W. (Eds.) (1970). *Percept genetic analysis*. Lund, Sweden: Gleerup.
Lawrence, T. R. (1993). Gathering in the sheep and goats: A meta-analysis of forced-choice sheep-goat ESP studies, 1947-1993. *Proceedings of Presented Papers: The Parapsychological Association 36th Annual Convention,* 75–86.
Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Götzsche, P. C., Ioannidis, J. P. A., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *BMJ, 339*, b2700. doi:10.1136/bmj.b2700.
Luke, D. P., Delanoy, D., & Sherwood, S. J. (2003). *Questionnaire of beliefs about luck*. Unpublished instrument, University of Northampton, UK.
*Luke, D. P., Delanoy, D., & Sherwood, S. J. (2008). Psi may look like luck: Perceived luckiness and beliefs about luck in relation to precognition. *Journal of the Society for Psychical Research, 72*, 193–207.
*Luke, D. P., & Morin, S. (2009). *Luck beliefs, PMIR, psi and the sheep-goat effect: A replication.* Abstracts of papers presented at the 33rd International Conference of the Society for Psychical Research, University of Nottingham (pp. 12–13).
*Luke, D. P., Roe, C. A., & Davison, J. (2008). Testing for forced-choice precognition using a hidden task: Two replications. *Journal of Parapsychology, 72*, 133–154.
Mangan, G. L. (1958). *A review of published research on the relationship of some personality variables to ESP scoring level*. New York, NY: Parapsychology Foundation.
McCarthy, D., & Schechter, E. I. (1986). Estimating effect size from critical ratios [Abstract]. In D. H. Weiner & D. I. Radin (Eds.), *Research in parapsychology 1985* (pp. 95–96). Metuchen, NJ: Scarecrow Press.
McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology, 52*, 81–90. doi:10.1037/0022-3514.52.1.81
Milton, J., & Wiseman, R. (1999). Does Psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin*, *125*, 387–391. doi:10.1037/0033-2909.125.4.387
Mousseau, M. C. (2003). Parapsychology: Science or pseudo-science. *Journal of Scientific Exploration*, *17*, 271–282.
Murray, A. L. (2011). The validity of the meta-analytic method in addressing the issue of psi replicability. *Journal of Parapsychology*, *75*, 261–277.
Nazimuddin, S. K. (2015). A study of individual differences in educational situations. *International Journal of Scientific Engineering and Research, 3*, 180–184.
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349,* aaac4716-1–aaac4716-8. doi:10.1126/science.aac4716
Palmer, J. (1978). Extrasensory perception: Research findings. In S. Krippner (Ed,), *Advances in parapsychological research 2* (pp. 59–243). New York, NY: Plenum. doi:10.1007/978-1-4615-9092-7_3

Palmer, J. (2003). ESP in the ganzfeld: Analysis of a debate. *Journal of Consciousness Studies*, *10*, 51–68.
*Palmer, J. (2009). Decision augmentation in a computer guessing task. J*ournal of Parapsychology, 73*, 119–136.
Palmer, J. (2016). Hansel's ghost: Resurrection of the experimenter fraud hypothesis in parapsychology [Editorial]. *Journal of Parapsychology*, *80*, 5–16.
Persinger, M. A., & Makarec, K. (1987). Temporal lobe epileptic signs and correlative behaviors displayed by normal populations. *Journal of General Psychology, 114,* 179–195. doi:10.1080/00221309.1987.9711068
Rhine, J. B., Pratt, J. G., Stuart, C. E., Smith, B. M., & Greenwood, J. A. (1940). *Extra-sensory perception after sixty years: A critical appraisal of the research in extra-sensory perception*. New York, NY: Henry Holt. doi:10.1037/13598-000
Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638–641. doi:10.1037/0033-2909.86.3.638
Rothstein, H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology*, *4*, 61–81. doi:10.1007/s11292-007-9046-9
Schmeidler, G. R. (1943). Predicting good and bad scores in a clairvoyance experiment. A preliminary report. *Journal of the American Society for Psychical Research*, *37*, 103–110.
*Schmeidler, G. R. (1964a). An experiment on precognitive clairvoyance. Part I. The main results. *Journal of Parapsychology*, *28*, 1–14.
*Schmeidler, G. R. (1964b). An experiment on precognitive clairvoyance, Part III. Precognition scores related to the subject's way of viewing time. *Journal of Parapsychology, 28*, 94–101.
*Schmeidler, G. R. (1964c). An experiment on precognitive clairvoyance: Part IV. Precognition scores related to creativity. *Journal of Parapsychology, 28*, 102–108.
*Schmeidler, G. R. (1964d). An experiment on precognitive clairvoyance: Part V. Precognition scores related to feelings of success. *Journal of Parapsychology, 28*, 109–125.
Schmidt, S., Schneider, R., Utts, J., & Walach, H. (2004). Distant intentionality and the feeling of being stared at: Two meta-analyses. *British Journal of Psychology, 95*, 235–247. doi:10.1348/000712604773952449
Smith, M. D. (1998). *Perception of one's own luck: The formation, maintenance and consequences of perceived luckiness* (Unpublished doctoral thesis). University of Hertfordshire, UK.
*Stanford, R. G. (1970). Extrasensory effects upon memory. *Journal of the American Society for Psychical Research*, *64*, 161–186.
Steinkamp, F. (1998). Psychology, psi, and the web: An exploratory study. *Journal of the American Society for Psychical Research, 92*, 256–278.
Steinkamp, F. (2005). Forced-choice ESP experiments: Their past and their future. In M. A. Thalbourne & L. Storm (Eds.), *Parapsychology in the 21st century: Essays on the future of psychical research* (pp. 124–163). Jefferson, NC: McFarland.
Steinkamp, F., Milton, J., & Morris, R. L. (1998). A meta-analysis of forced-choice experiments comparing clairvoyance and precognition. *Journal of Parapsychology*, *62*, 193–218.
*Storm, L. (2002). A parapsychological investigation of the I Ching: Seeking psi in an ancient Chinese system of divination. *Australian Journal of Parapsychology, 2*, 44–62.
Storm, L. (2006). Technical Paper No. 11: Meta-analysis in parapsychology: I. The ganzfeld domain. *Australian Journal of Parapsychology*, *6*, 35–53.
*Storm, L. (2008). Investigations of the I Ching: I. Relationships between psi and time perspective, paranormal belief and meaningfulness. *Australian Journal of Parapsychology*, *8*, 103–127.
*Storm, L., & Thalbourne, M. A. (1998-1999). The transliminal connection between personality and paranormal effects in an experiment with the I Ching. *European Journal of Parapsychology, 14*, 100–124.
Storm, L., Tressoldi, P. E., & Di Risio, L. (2012). Meta-analysis of ESP studies, 1987-2010: Assessing the success of the forced-choice design in parapsychology. *Journal of Parapsychology*, *76*, 243–273.
Tart, C. T. (1983). Information acquisition rates in forced-choice ESP experiments: Precognition does not work as well as present-time ESP. *Journal of the American Society for Psychical Research, 77*, 293–310.
*Thalbourne, M. (1996). An attempt to predict precognition scores using transliminality-relevant variables. *Journal of the Society of Psychical Research, 62*, 129–140.
Thalbourne, M. A. (1998). Transliminality: Further correlates and a short measure. *Journal of the American Society for Psychical Research, 92*, 402–419.
Thalbourne, M. A., & Delin, P. S. (1993). A new instrument for measuring the sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychical Research, 59*, 172–186.
Thalbourne, M. A., & Delin, P. S (1994) A common thread underlying belief in the paranormal, mystical experience

and psychopathology. *Journal of Parapsychology, 58*, 3–38.
*Thalbourne, M. A., & Storm, L. (2014). A further study of psychopraxia using the "I Ching." *Australian Journal of Parapsychology*, *14*, 115–142.
Torrance, E. P. (1963). *Preliminary Manual for Personal-Social Motivation Inventory*. Mineapolis, MN: University of Minnesota.
Utts, J. (1991). Replication and meta-analysis in parapsychology. *Statistical Science, 6*, 363–378.
Utts, J. (2015). What constitutes replication in parapsychology?. In E. C. May, & S. B. Marwaha (Eds.), *Extrasensory perception: Support, skepticism, and science* (Vol. 1). Santa Barbara, CA: ABC-CLIO.
Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics, 35*, 215–247. doi:10.3102/1076998609346961
*Vaughn, A., & Houck, J. (1993). A "success" test of precognition and attitude toward the future. *Journal of the American Society for Psychical Research*, *59*, 259–268.
Watt, C. (2005). Parapsychology's contribution to psychology: A view from the front line. *Journal of Parapsychology*, *69*, 215–231.
Watt, C. (2016). A prospective meta-analysis of pre-registered ganzfeld ESP studies. KPU Study Registry. Retrieved from http://www.koestler-parapsychology.psy.ed.ac.uk/Documents/KPU_Registry_1024.pdf
Watt, C., & Kennedy, J. E. (2015). Lessons from the first two years of operating a study registry. *Frontiers in psychology*, *6*, 173. doi:10.3389/fpsyg.2015.00173
Watt, C., & Kennedy, J. E. (2016). Options for prospective meta-analysis and introduction of registration-based prospective meta-analysis. *Frontiers in Psychology, 7*, 2030. doi:10.3389/fpsyg.2016.02030
*Winkelman, M. (1981). The effect of formal education on extrasensory abilities: The Ozolco study. *Journal of Parapsychology*, *45*, 321–336.
Wiseman, R., & Greening, E. (2002). The Mind Machine: A mass participation experiment into the possible existence of extrasensory perception. *British Journal of Psychology*, *93*, 487–499. doi:10.1348/000712602761381367

Meta-Analysis Papers Not Cited

Haraldsson, E., & Johnson, M. (1979). ESP and the Defense Mechanism Test (DMT) Icelandic study No. III: A case of the experimenter effect? *European Journal of Parapsychology, 3*, 11–20.
Haraldsson, E. (1978). ESP and the defense mechanism test (DMT): A further validation. *European Journal of Parapsychology*, *2*, 104–114.
Haraldsson, E., & Johnson, M. (1986). The Defense Mechanism Test (DMT) as a predictor of ESP performance: Icelandic studies VI and VII [Abstract]. In D. H. Weiner & D. I. Radin (Eds.), *Research in parapsychology 1985* (pp. 43–44). Metuchen, NJ: Scarecrow Press.
Honorton, C. (1967). Creativity and precognition scoring level. *Journal of Parapsychology*, *31*, 29–42.
Johnson, M. (1969). Attitude and target differences in a group precognition test. *Journal of Parapsychology*, 33, 324–325.
Johnson, M., & Haraldsson, E. (1984). The Defense Mechanism Test as a predictor of ESP scores: Icelandic studies IV and V. *Journal of Parapsychology*, *48*, 185–200.
Nielsen, W. (1970). Relationships between precognition scoring level and mood. *Journal of Parapsychology*, *34*, 93–116.
Ryzl, M. (1968a). Precognition scoring and attitude toward ESP. *Journal of Parapsychology*, *32*, 1–8.
Ryzl, M. (1968b). Precognition scoring and attitude. *Journal of Parapsychology*, *32*, 183–189.
Schwartz, S. A., & De Mattei, R. J. (1983). The Mobius Psi-Q Test: Preliminary findings [Abstract]. In R. A. White & J. Solfvin (Eds.), *Research in parapsychology 1984* (pp. 103–105). Metuchen, NJ: Scarecrow Press.
Smith, M. D., Wiseman, R., Machin, D., Harris, P., & Joiner, R. (1997). Luckiness, competition and performance on a psi task. *Journal of Parapsychology, 61*, 33–43.
Storm, L. (2006). A parapsychological investigation of the I Ching: The relationship between psi, intuition, and time perspective. *Journal of Parapsychology*, *70*, 121–141.
Storm, L., & Thalbourne, M. A. (2001). Studies of the I Ching: I. A replication. *Journal of Parapsychology, 65*, 105–124.
Thalbourne, M., Beloff, I., & Delanoy, D. (1982). A test for the "extraverted sheep versus introverted goats" hypothesis [Abstract]. In W. G. Roll, R. L. Morris, & R. A. White (Eds.), *Research in parapsychology 1981* (pp. 155–156). Metuchen, NJ: Scarecrow Press.
Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An agenda for purely

confirmatory research. *Perspectives on Psychological Science*, *7*, 632–638. doi:10.1177/1745691612463078
Wilson, S., & Hamlin, I. (2007). Implicit learning in a card prediction task. *European Journal of Parapsychology, 22*, 3–29.

Acknowledgements

The authors would like to thank Caroline Watt, Gergö Hadlaczky, and Jan Dalkvist for their comments on an earlier draft of this paper. We would also like to thank John Palmer and the two anonymous referees for their constructive feedback.

Victoria University of Wellington
Kelburn Parade
PO Box 600
Wellington 6012, New Zealand
Marco.Zdrenka@vuw.ac.nz

Appendix

Table A1
Number of Studies for Each Individual Difference Measure That Reported Nonsignificance Without Providing Any Other Statistical Information

Individual Difference Measure	Number of non-significant studies estimated as $r = .00$
Belief in Psi	5
Extraversion	3
Neuroticism	2
Agreeableness	2
Spontaneous Experiences	2
Openness to Experience	1
Dream Recall	1
Conscientiousness	1
Temporal Lobe Dysfunction	1
Cerebral Lateralization	1

Table A2
68% and 80% Confidence Intervals for All Individual Difference Measures Meta-Analysed

Individual Difference Measure	68% Confidence Intervals	80% Confidence Intervals
Chance Belief	.03, .25	-.01, .28
Luck Belief	.07, .20	.05, .21
Belief in Psi	.09, .16	.07, .18
Perceptual Defensiveness	.06, .19	.04, .20
Providence Belief	.05, .20	.02, .22
Openness to Experience	.07, .17	.06, .19
Extraversion	.05, .11	.04, .12
Perceived Luckiness	-.03, .18	-.06, .21
Dream Recall	.01, .13	-.004, .14
Intelligence	-.19, .07	-.23, .10
Emotional Reactivity	-.11, .23	-.16, .28
Conscientiousness	-.02, .13	-.04, .15
Neuroticism	-.01, .12	-.03, .14
Religiosity	-.15, .05	-.18, .07
Future-Oriented	-.04, .15	-.07, .17
Creativity	-.02, .11	-.04, .13
Fortune Belief	-.03, .10	-.05, .12
Agreeableness	-.04, .08	-.05, .10
Pro Attitude	-.07, .03	-.08, .05
Time Belief as Dynamic	-.03, -.01	-.03, -.01
Transliminality	-.06, .08	-.08, .10
Spontaneous Experiences	-.04, .03	-.05, .04
Temporal Lobe Dysfunction	-.10, .09	-.13, .12