Psychological Bulletin
2010, Vol. 136, No. 4, 491–494


0033-2909/10/$12.00 DOI: 10.1037/a0019840

# A Meta-Analysis With Nothing to Hide: Reply to Hyman (2010)

Lance Storm
University of Adelaide

Patrizio E. Tressoldi and Lorenzo Di Risio
Università di Padova

In our article (Storm, Tressoldi, & Di Risio, 2010), we claimed that the ganzfeld experimental design has proved to be consistent and reliable. However, Hyman (2010) argues that the overall evidence for psi is, in fact, contradictory and elusive. We present a case for psi research that undermines Hyman's argument. First, we give examples from parapsychologists who do not outrightly dismiss psi, despite appearances, but actually support it. Second, we claim that Hyman does not tell the full story about the ganzfeld meta-analytic findings and thus presents a one-sided account. Third, we argue that our meta-analysis has followed standard procedures, that we have not broken any rules but have found a communications anomaly, often referred to as psi. Though we may be in agreement that the evidence is largely statistical, the evidence suggests that concealed targets are actually identified rather than guessed. We argue that further research is necessary.

*Keywords:* ESP, free response, ganzfeld, meta-analysis, psi

## On the Words Used to Describe Evidence for Psi

In our article (Storm, Tressoldi, & Di Risio, 2010), we stated that "the ganzfeld is one of the most consistent and reliable experimental paradigms in parapsychology" (p. 479). We believe that the empirical evidence supports our claim, but in response to our statement, Hyman (2010) makes a case from an apparently broader "context" and alleges that various authors (i.e., Atmanspacher & Jahn, 2003; Bierman, 2001; Kennedy, 2001, 2003; and Lucadou, 2001) have conversely claimed that the evidence for psi is "inconsistent, unreliable, contradictory, and elusive" (p. 486). We point out that we were not making claims about psi but were commenting on an experimental design. On that note, Hyman states it is "unclear" whether our claims for the ganzfeld "should impress the scientific community" (Hyman, 2010, p. 486), but we are on record for saying that "parapsychologists may still have some way to go to convince skeptics" (Storm et al., 2010, p. 480). Although we may agree with Hyman to some degree about how the ganzfeld may be perceived, we wish to critique Hyman's more general allegation against the evidence for psi because we believe it is not entirely warranted, especially given the literature he cites.

We disagree with Hyman's (2010) approach of citing an arbitrarily representative group of four or five parapsychologists and, in some cases, putting words in their mouths. For example, Bierman (2001) used only one of the four above-mentioned adjectives (i.e., *elusive*) and did so only in the context of an explanation by Lucadou (2001) of the "elusive character of laboratory experiments" (Lucadou, 2001, p. 10), not psi per se. More importantly, Bierman refers to the "cumulating evidence" of psi and adds that the "evidence is very strong" (p. 269). Furthermore, Lucadou states that "individual researchers (including me) are personally convinced that psi exists" (p. 7). Importantly, neither Bierman nor Lucadou regard psi as the result of "statistical flukes or errors" (Bierman, 2001, p. 269) or "flaws or artifacts" (Lucadou, 2001, p. 15).

Of course, some researchers, including Kennedy (2001), might feel the aforementioned adjectives are appropriate, but Kennedy (2003) is effectively impartial: "psi is sometimes impressive and reliable" (p. 53). And Atmanspacher and Jahn (2003) are more than open-minded:

> A huge body of reliable data is now available. Meta-analytic evaluation of these sources remains difficult for a number of reasons (see Radin & Nelson, 1989; and Ehm, 2003), but is indicative of widespread and recurrent anomalous effects which must not be ignored, but are yet to be explained. (p. 261)

We would add additional quotes. As Hyman notes, Dean Radin (a distinguished longtime contributor to the parapsychological debate) describes psi effects as "*consistent*" (Radin, 1997, p. 58, emphasis in the original; cited in Hyman, 2010, p. 489). And Honorton and Ferrari (1989), in their meta-analysis of precognition forced-choice experiments, refer to "robust" and highly significant results across a time span of more than 50 years without evidence of a decline. This kind of exchange could go on ad infinitum, but little progress would be made.

## On Meta-Analysis

Hyman (2010) undermines the objectives and gains made from meta-analysis, referring to it as having "serious limitations . . . as a method for confirming hypotheses and for establishing the replicability of experiments" (p. 486). In fact, meta-analyses in all scientific fields are conducted to *clarify* findings in a collection of related studies for which there may be some ambiguity at the unit level of the individual study. Starting with a research hypothesis (i.e., that there is an effect), investigators seek to *confirm* (i.e., find

Lance Storm, School of Psychology, University of Adelaide, Adelaide, Australia; Patrizio E. Tressoldi and Lorenzo Di Risio, Dipartimento di Psicologia Generale, Università di Padova, Padua, Italy.

Correspondence concerning this article should be addressed to Lance Storm, School of Psychology, University of Adelaide, Adelaide SA 5005, Australia. E-mail: lance.storm@adelaide.edu.au

support) for that hypothesis. From these findings, useful policies may be subsequently adopted (see, e.g., Cook et al., 1992; Glass, McGaw, & Smith, 1981).

Hyman (2010) argues that it is "prospective evidence for replicability, not the retrospective kind that emerges from meta-analyses of previous experiments, that is required for scientific acceptability" (p. 486). We find that Hyman ironically makes too big a case for the ganzfeld, and describes the upshot of our article as a bid for a standard ganzfeld recipe that parapsychologists might follow. But we never intended to prescribe or schematize the perfect psi experiment because we also argued that nonganzfeld "noise reduction techniques elicit statistical evidence for a communications anomaly" on par with the ganzfeld (Storm et al., 2010, p. 480). Nevertheless, Hyman appeals to the Autoganzfeld II study by Broughton and Alexander (1997), where the aforementioned ganzfeld "recipe" had "failed" because the hit rate was an overall 25.8% for 209 trials. In fact, the "emotionally close series" produced a significant hit rate of 37.3%, and the "general series" produced a comparable hit rate of 37.5%.

Before Hyman (2010) discusses the Autoganzfeld II study, he refers to the Autoganzfeld I study, contending that the "significant hit rate obtained in the autoganzfeld [I] experiments was due entirely to the hit rate of 37% for the dynamic targets" (pp. 486–487). We note, however, that Bem and Honorton (1994) actually "predicted" (p. 12) a significantly higher hit rate for dynamic targets (a significant 37%) as "suggested by the earlier [Honorton et al., 1990] meta-analysis" (p. 12)—a proof that meta-analysis can provide "prospective evidence for replicability" (Hyman, 2010, p. 486). It must also be remembered that the nonsignificant 27% for the "static target" study set is still above mean chance expectation, and every data set that is above chance makes a contribution to the overall ganzfeld trend. As for participant type, Bem and Honorton noted that "88% of the participants reported personal experiences suggestive of psi, and 80% had some training in meditation or other . . . mental disciplines" (p. 13). We note that the samples in Broughton and Alexander's (1997) study were mostly too small for them to yield significant correlations, but those participants who reported "psi experiences" produced a higher mean hit rate (26.8%) than those who did not (23.1%), although the effect was not in the direction hypothesized for those who practiced "mental disciplines" (p. 219).

Hyman (2010) also mentions Bierman's (2001) failed attempt at ganzfeld replication, but out of four studies conducted by Bierman, one was significant and another approached significance (see Bierman, 1995; Bierman, Bosga, Gerding, & Wezelman, 1993). And because Hyman makes an issue of replication, why does he not mention four other significant replications conducted during that relevant decade (Dalton, 1994; Kanthamani & Broughton, 1994, Study 4; Morris, Cunningham, McAlpine, & Taylor, 1993, Study I; Morris, Dalton, Delanoy, & Watt, 1995)?

The issue of concern here is whether one or two isolated studies provide sufficient evidence against an experimental design. We know that psi effects are generally weak, as are many psychological effects, so we do not expect replication on demand in any given study. Again, that is why we conduct meta-analyses. If the meta-analyses tell us anything, it is that nothing is absolute.

## On the Capricious Nature of Psi

Hyman itemizes "problems" that Kennedy (2003) has identified as indicating the "capricious" nature of psi, including "psi missing and negative reliability, the shift from intended effects to unintended secondary effects, erosion of evidence and decline effects over time, the inverse correlation of effect size with improved methodology, and the lack of practical applications of psi" (Hyman, 2010, p. 487). For the ganzfeld, we have already argued in our article that the "erosion of evidence," "negative reliability," and "the inverse correlation of effect size with improved methodology" are in dispute. The other effects ("psi missing" and "secondary effects") have been well documented and independently tested by researchers. For example, psi missing can be taken as a mirror effect more commonly produced by nonbelievers, just as psi hitting is more common among believers (Lawrence, 1993).

As for practical applications of psi, this is largely unexplored territory (especially in the laboratory), but for many years we have been aware of psychic individuals who are, or have been, gainfully employed in psi practice and produce useful or profitable results (see Schouten, 1993; Schwartz, 1983, 2000, 2005). In his comprehensive review, Schouten (1993) found that "psychics might occasionally have impressions which are difficult to explain and which could be considered paranormal" (p. 387), also commenting that "complementary medicine deserves a place in the health care system" (p. 399). Schwartz (1983, 2000, 2005) has gone to great lengths to demonstrate the uses of psi in archaeology and anthropology, and this approach extends to crime solving and detection—areas that have already seen applications (Schouten, 1994).

Claiming that our meta-analysis "masks rather than uncovers the actual situation" (p. 487), Hyman (2010) alleges that we have "manufactured" consistency. We state that removing outliers to create homogeneous databases and/or to merge databases is standard practice in meta-analysis. We were transparent in our procedures and findings, applied standard rules, and did not invent any untoward techniques (see, e.g., Rosenthal, 1984). If the "original populations of experiments are heterogeneous" (Hyman, 2010, p. 487), then so be it—surely it is a better practice to err on the side of caution than mistakenly inflate or deflate the size of an effect. As for statistical testing of hypotheses, it is a truism that statistical proof of the null is never possible.

On whether the Milton and Wiseman (1999) database is an "outlier," we meant it to be counted as an exception because our population estimates of $z$ score and effect size values included zero for that database, and only that database. Even so, we put the psi hypothesis severely to test in finding that such a low-yield database (effect size = .013) could still be legitimately merged with other databases without adversely affecting the overall significance of increasingly larger databases.

We also stress that Hyman's (2010) claim is simply untrue that a "rebound effect implies that the effect sizes of the most recent studies are *significantly* [emphasis added] higher than those of the immediately preceding studies" (p. 487). Hyman's use of the word "significantly" suggests that heterogeneous databases were combined, which is not the case. The combined 108 studies is the result of consistently finding nonsignificant differences between databases.

Finally, "using the exact binomial to calculate a combined effect size and test it for significance" (Hyman, 2010, p. 487) is entirely acceptable. As we reported in our article (Storm et al., 2010, p. 473), statistician Jessica Utts (personal communication, December 11, 2009) endorsed and used the exact binomial test on trial counts in the Milton and Wiseman (1999) database and found a significant hit rate of 27%. We had used standard meta-analytic procedures, and there was no "mistake" (Hyman, 2010, p. 487) in our actions. On the contrary, we presented a truer picture of the variance in effects by including rather than excluding the Milton and Wiseman database on the basis of statistical test results.

Hyman (2010) states that "combining of effect sizes from different studies makes sense only if one can show that the separate effect sizes are conceptually coherent—that they all can be attributed to the same underlying cause," adding that "parapsychologists admit that they have no positive theory for deciding which departures from chance, if any, reflect the presence of a single entity called psi" (p. 487). We acknowledge that "no positive theory" yet exists to explain psi, but our article was neither a theoretical exercise nor a comparison of psi theories (for which there are many; see Stokes, 1987). We took a purely empirical approach by presenting an update. We believe that we confirmed that psi effects can be found in a larger data set where any given single study on its own may be inconclusive. Our action is no less inappropriate than Galileo's demonstration of gravitational effects in pre-Newtonian times without it being necessary for Galileo to have a "positive theory" that describes why such effects should exist.

Hyman (2010) refers to the "surprising number of problems and flaws" (p. 488) in Honorton's (1985) database, as documented in his (Hyman, 1985) critique, but the issues he raised have been disputed by numerous respondents including Harris and Rosenthal (1988), Palmer (1986), and Saunders (1985). (Space does not permit elaboration here.) On one particular issue, Hyman argues that the same database by Honorton was heterogeneous, comprising a mixture of studies from two populations: (a) "four experimenters who consistently contributed experiments with above-chance results [44%]" and (b) "several other experimenters who consistently obtained results consistent with chance [26%]" (p. 488). A worthy point is demonstrated here, but to put it into perspective, it contrasts with our finding that significant experimenter/laboratory differences were not found (see Storm et al., 2010, p. 477). This finding is a replication (see Bem & Honorton, 1994; Honorton et al., 1990).

## On Communication Anomalies

Hyman (2010) argues that our claim for "communication anomalies" only "begs the question" (p. 488) because the nature of the anomaly has not been specified. We agree that the effect is a statistical anomaly first and foremost, but there is clearly some degree of unexplained identification of targets above chance expectation, which surely gives cause for further investigation. Science needs imaginative researchers who can find useful applications for this effect. The alternative is to ignore the effects and hope, as many scientists do, that these alleged effects will go away as people wise up and the small community of parapsychologists loses interest, so that the issue will never be resolved.

It is difficult to discern what Hyman (2010) intends for parapsychology. We can only say that a true and proper attitude to psi phenomena is to take them seriously until proven otherwise, which surely requires a consistent and sustained effort. Certainly, given the little time devoted to parapsychology compared with other disciplines, psychical research is still young. We thus believe that parapsychology is unfairly dismissed by skeptics who state that "despite over a century of failing to come up with even one replicable experiment, parapsychological claims are still with us" (Hyman, 2010, p. 489). Across this century of work, limited resources were available to the field, and replications have been discounted. Proffering a few studies failing to reveal replication does not stand up against the overall findings from the greater weight of many dozens of studies in the meta-analyses. Independent of the meta-analyses, repeatable psi effects in single studies have been found, even by skeptics like Stanley Jeffers (Freedman, Jeffers, Saeger, Binns, & Black, 2003).

Hyman (2010) states that we believe "parapsychological claims are dismissed for reasons other than the adequacy of their evidence" (p. 488); in large part, this is true. Mainstream academia cannot see a place for psi. We argue that their interest would quickly change if we could furnish them with what they deem to be evidence, but because many academics have "already decided the issue," to use Hansen's (1991, p. 202) words, minds are closed. It is not only the spirit of academic freedom that is violated or at least curbed but, more broadly, the scientific enterprise.

Hyman (2010) closes by stating the following: "Required, of course, are demonstrations that the claimed evidence can be prospectively obtained by independent investigators, given appropriately designed experiments with adequate power" (p. 489). We remind readers that we tested laboratories and investigators and found that no single experimenter or laboratory group produced effect sizes significantly different from any other. As noted above, this finding is a replication.

## Conclusion

Hyman (2010) is convinced of parapsychology's "persistent inconsistency" (p. 489), but we maintain that psi is anything but inconsistent. We argue that the meta-analytic results do count for something. It is of paramount importance not to dismiss statistical anomalies as nothing more than numerical oddities; we would hope that on the strength of the peculiarity of the findings in and of themselves, the scientifically inclined will be sufficiently intrigued by the curious nature of psi to want to find out for themselves what that anomaly might actually be.

On that point, Hyman (2010) claims the following:

> Parapsychology will succeed in its quest to demonstrate its communications anomaly only when it can generate specific hypotheses that predict patterns of outcomes that are consistent, lawful, and independently replicable by parapsychologists and others. So far, careful assessment of the parapsychological literature does not justify optimism on this matter. (p. 490)

In response, we believe had addressed these issues, and argue for a shift in mind-set: Instead of parapsychologists' giving the null hypothesis a chance (Alcock, 2003), skeptics should give the alternative hypothesis a chance. In spite of the relatively limited pool of literature, we argue that consistency has been demonstrated in the data and that there is good evidence of replication by a range of investigators. We maintain that parapsychology is a struggling discipline despite its age, and we can only imagine where science might be if parapsychology had been given the same due attention that other disciplines have received over the same period.

## References

Alcock, J. E. (2003). Give the null hypothesis a chance: Reasons to remain doubtful about the existence of psi. In J. E. Alcock, J. E. Burns, & A. Freeman (Eds.), *Psi wars: Getting to grips with the paranormal* (pp. 29–50). Charlottesville, VA: Imprint Academic.

Atmanspacher, H., & Jahn, R. G. (2003). Problems of reproducibility in complex mind–matter systems. *Journal of Scientific Exploration, 17,* 243–270.

Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin, 115*(1), 4–18. doi:10.1037/0033-2909.115.1.4

Bierman, D. J. (1995). The Amsterdam ganzfeld series III & IV: Target clip emotionality, effect sizes and openness. In *Proceedings of the 38th Annual Convention of the Parapsychological Association* (pp. 27–37). Durham, NC: Parapsychological Association.

Bierman, D. J. (2001). On the nature of anomalous phenomena: Another reality between the world of subjective consciousness and the objective world of physics? In P. Van Loocke (Ed.), *The physical nature of consciousness* (pp. 269–292). New York, NY: Benjamins.

Bierman, D. J., Bosga, D. J., Gerding, H., & Wezelman, R. (1993). Anomalous information access in the ganzfeld: Utrecht—Novice series I and II. In *Proceedings of the 36th Annual Convention of the Parapsychological Association* (pp. 192–204). Durham, NC: Parapsychological Association.

Broughton, R. S., & Alexander, C. H. (1997). Autoganzfeld II: An attempted replication of the PRL ganzfeld research. *Journal of Parapsychology, 61,* 209–226.

Cook, T. D., Cooper, H., Cordray, D. S., Hartmann, H., Hedges, L. V., Light, R. J., . . . Mosteller, F. (1992). *Meta-analysis for explanation: A casebook.* New York, NY: Russell Sage Foundation.

Dalton, K. (1994). A report on informal ganzfeld trials and comparison of receiver/sender sex pairing: Avoiding the file drawer. In D. J. Bierman (Ed.), *The Parapsychological Association 37th Annual Convention: Proceedings of presented papers* (pp. 104–113). Durham, NC: Parapsychological Association.

Ehm, W. (2003). Pattern count statistics for the analysis of time series in mind–matter studies. *Journal of Scientific Exploration, 17,* 497–520.

Freedman, M., Jeffers, S., Saeger, K., Binns, M., & Black, S. (2003). Effects of frontal lobe lesions on intentionality and random physical phenomena. *Journal of Scientific Exploration, 17,* 651–668.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* London, England: Sage.

Hansen, G. P. (1991). The elusive agenda: Dissuading as debunking in Ray Hyman's "The Elusive Quarry." *Journal of the American Society for Psychical Research, 85,* 193–203.

Harris, M. J., & Rosenthal, R. (1988). *Postscript to "Human performance research: An overview."* Washington, DC: National Academies Press.

Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology, 49,* 51–91.

Honorton, C., Berger, R. E., Varvoglis, M. P., Quant, M., Derr, P., Schechter, E. I., & Ferrari, D. C. (1990). Psi communication in the ganzfeld: Experiments with an automated testing system and a comparison with a meta-analysis of earlier studies. *Journal of Parapsychology, 54,* 99–139.

Honorton, C., & Ferrari, D. C. (1989). "Future telling": A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology, 53,* 281–308.

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*(4), 486–490. doi:10.1037/a0019676

Kanthamani, H., & Broughton, R. S. (1994). Institute for Parapsychology ganzfeld–ESP experiments: The manual series. In D. J. Bierman (Ed.), *The Parapsychological Association 37th Annual Convention: Proceedings of presented papers* (pp. 182–189). Durham, NC: Parapsychological Association.

Kennedy, J. E. (2001). Why is psi so elusive? A review and proposed model. *Journal of Parapsychology, 65,* 219–246.

Kennedy, J. E. (2003). The capricious, actively evasive, unsustainable nature of psi: A summary and hypotheses. *Journal of Parapsychology, 67,* 53–74.

Lawrence, T. R. (1993). Gathering in the sheep and goats: A meta-analysis of forced-choice sheep–goat ESP studies, 1947–1993. In *Proceedings of the 36th Annual Convention of the Parapsychological Association* (pp. 75–86). Durham, NC: Parapsychological Association.

Lucadou, W. V. (2001). Hans in luck: The currency of evidence in parapsychology. *Journal of Parapsychology, 65,* 3–16.

Milton, J., & Wiseman, R. (1999). Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin, 125*(4), 387–391. doi:10.1037/0033-2909.125.4.387

Morris, R. L., Cunningham, S., McAlpine, S., & Taylor, R. (1993). Toward replication and extension of autoganzfeld results. In *Proceedings of the 36th Annual Convention of the Parapsychological Association* (pp. 177–191). Durham, NC: Parapsychological Association.

Morris, R. L., Dalton, K., Delanoy, D. L., & Watt, C. (1995). Comparison of the sender/no sender condition in the ganzfeld. In *Proceedings of the 38th Annual Convention of the Parapsychological Association* (pp. 244–259). Durham, NC: Parapsychological Association.

Palmer, J. (1986). Comments on the "joint communiqué." *Journal of Parapsychology, 50,* 377–381.

Radin, D. (1997). *The conscious universe: The scientific truth of psychic phenomena.* San Francisco, CA: HarperEdge.

Radin, D. I., & Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics, 19,* 1499–1514.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

Saunders, D. R. (1985). On Hyman's factor analysis. *Journal of Parapsychology, 49,* 86–88.

Schouten, S. A. (1993). Applied parapsychology studies of psychics and healers. *Journal of Scientific Exploration, 7,* 375–401.

Schouten, S. A. (1994). An overview of quantitatively evaluated studies with mediums and psychics. *Journal of the American Society for Psychical Research, 88,* 221–254.

Schwartz, S. (1983). Preliminary report on a prototype applied parapsychological methodology for utilization in archaeology, with a case report. In W. G. Roll, R. L. Morris, & R. White (Eds.). *Research in parapsychology 1981* (pp. 25–27). Metuchen, NJ: Scarecrow Press.

Schwartz, S. (2000). *The Alexandria Project.* New York, NY: Authors Guild.

Schwartz, S. (2005). *The secret vaults of time.* Charlottesville, VA: Hampton Roads.

Stokes, D. M. (1987). Theoretical parapsychology. In S. Krippner (Ed.), *Advances in parapsychological research* (Vol. 5, pp. 77–189). Jefferson, NC: McFarland.

Storm, L., Tressoldi, P. E., & Di Risio, L. (2010). Meta-analysis of free-response studies, 1992–2008: Assessing the noise reduction model in parapsychology. *Psychological Bulletin, 136*(4), 471–485. doi: 10.1037/a0019457

Received March 31, 2010
Revision received March 31, 2010
Accepted April 1, 2010 ■