KPU Registry ID Number: 1024 Date revised: 22nd Nov 2017
(This revised version was submitted prior to the registration of any studies that will be included)

Study Registration For the KPU Study Registry

The registration information for the study is given below.

1. The title or name of the experiment (for listing the experiment in the registry).

A prospective meta-analysis of pre-registered Ganzfeld ESP studies

2. The name, affiliation, and email address for the lead experimenter(s) for the study.

Professor Caroline Watt
University of Edinburgh
Caroline.Watt@ed.ac.uk

3. A short description or abstract of the purpose and design of the experiment.

A prospective meta-analysis of pre-registered ganzfeld ESP studies, with the inclusion criteria and outcome measures planned prior to the conduct of any studies to be included. The definition of ganzfeld, inclusion criteria, and planned analyses are based on the observations and recommendations of Bem, Palmer & Broughton (2001), Storm, Tressoldi & di Risio (2010), Baptista, Derakhshani & Tressoldi (2015), Higgins, Deeks & Altman (2011), and following consultation with psi proponents and psi skeptical researchers with expertise in ganzfeld, psi research, and statistical methodology.

4. A statement or list of the specific hypothesis or hypotheses being tested, and whether each hypothesis is confirmatory or exploratory. (confirm/explore guidance)

Confirmatory: The primary hypothesis is that evidence consistent with the operation of ESP will be found and the null hypothesis that the overall results are due only to chance will be rejected. Two analyses will be employed to test this hypothesis. One will be conducted on **telepathy** studies using the subset of participants who have been identified as meeting pre-specified criteria. The other will be conducted on the full dataset of studies including both selected and unselected participants and GESP (telepathy, clairvoyance or precognition) methods. An adjustment to alpha level for multiple analyses will be applied consistent with the Cochrane Collaboration guidance on systematic review (Higgins, Deeks & Altman, 2011) and as described in section 7 below. Studies must meet the criteria specified in section 7 below to be included.

Exploratory: Studies will be classified according to whether they do or do not have procedures that make it difficult for any one person to commit undetected fraud or to make undetected unintentional data alterations. The results for studies with and without these procedures will be compared.

Exploratory: If the null hypothesis is rejected and sufficient relevant data are available, factors that have been reported to be associated with success in previous studies will be evaluated[1]. These factors are: artistic, prior psi experience or training, prior psi belief, and practice of a mental discipline. However, these evaluations may be limited due to insufficient relevant data.

5. The planned number of participants and the number of trials per participant.

There are two confirmatory analyses. The first is for the full set of trials: for this analysis registered GESP studies will be added until a total of at least 921 trials have been included. The second confirmatory analysis is for the sub-set of telepathy studies with participants selected for artistic ability: for this analysis registered studies will be added until at least 234 trials have been included. The power analysis can be seen in section 8 below. Final analysis and final conclusions will be based on the planned sample size, whenever that is reached. If the planned sample size for either confirmatory hypothesis is reached then formal results will be announced for that hypothesis. Data collection for the other incomplete confirmatory hypothesis will continue until the required sample size is reached and the final results will then be formally announced for that hypothesis. Participants may provide more than one trial. The unit of analysis is the trial, not the participant.

6. A statement that the registration is submitted prior to testing the first participant, or indicating the number of participants tested when the registration (or revision to the registration) was submitted.

This meta-analysis is registered before the registration of any of the studies that will be included. Each included study must be registered before data collection begins for the study.

[1] This is an exploratory analysis rather than confirmatory because the primary purpose of this prospective meta-analysis is to evaluate the ESP hypothesis. Appropriate power analysis and study designs needed for confirmatory analyses of the factors that previous studies have suggested may be associated with better performance on the ESP task are not included in this meta-analysis plan. The evaluation of factors such as these as moderating variables in a meta-analysis is *synthesis-generated evidence* that is correlational analysis of observational data and is very susceptible to confounding. These analyses are primarily of value for motivating experiments with appropriate designs, including randomization to neutralize potential confounding factors. Such experiments provide *study-generated evidence* that is much more convincing and should be evaluated in future meta-analyses focused on a specific effect (Cooper & Hedges, 2009; Kennedy, 2013).

The following additional information is needed for studies that include confirmatory analyses:

7. Specification of all analysis decisions that could affect the confirmatory results, including: the specific statistical test for each confirmatory hypothesis, whether the test is one-sided or two-sided, the criterion for acceptable evidence, any transformations or adjustments to the data, any criteria for excluding or deleting data, and any corrections for multiple analyses. Checklists and examples for registering classical analyses, permutation and bootstrap analyses, Bayesian analyses, and classification analyses are provided in the statistics registration document. (This information can be included in section 4 above for simple experiments.)

Planned Analyses

The primary confirmatory hypothesis that the results support the ESP hypothesis rather than being due only to chance will be assessed using two tests of exact binomial probability on the pooled counts of direct hits and misses. The first test will be conducted on the pooled count of direct hits and misses from all participants, including the artistic participants, from studies using a telepathy, clairvoyance or precognition design, and with the alpha value for significance being $p \leq .04$ one-tailed. The second test will be conducted on the subset of telepathy studies that have all participants selected as artistic (e.g., visual artists, musicians, actors, and writers) plus artistic participants from other studies, in which the registration for the experiment included identification of an artistic subgroup. This is designated the 'artistic' subset. The alpha value for significance for this test will be $p \leq .01$ one-tailed. This allocation of alpha values applies the Bonferroni correction for multiple analyses, because two tests are being used to assess the experimental hypothesis. An unequal allocation (summing to .05) is planned because previous research findings indicate a much larger effect size is expected for artistic participants (Baptista, Derakhshani & Tressoldi, 2015; Dalton, 1997).

If either analysis meets the criterion for significance, then it will be concluded that evidence in support of the ESP hypothesis has been obtained. The assignment of studies to the artistic participants sub-group, and to telepathy, clairvoyance or precognition design will be made at the time of study registration, before study results are known. Given that the powers for the analyses are .95 for the artistic subset and .95 for the pooled total (see the power analysis below), a nonsignificant outcome for an analysis is evidence that the predicted effect size specified in the power analysis is false for this meta-analysis. Given that the effect sizes in the power analyses (30.0% hit rate for the pooled data and 37.0% hit rate for the artistic subset) were from the low end of the confidence intervals from previous studies, a nonsignificant outcome for an analysis is evidence that the minimum predicted effect size based on previous research is not true for this meta-analysis.

The specific methods for analyzing the exploratory hypotheses are not pre-specified. If sufficient relevant data are available, effect sizes and confidence intervals will be

estimated for the factors identified in section 4 above. Heterogeneity and random effects will be evaluated as appropriate for the available data.

Inclusion Criteria and Study Design Recommendations

The inclusion criteria are informed by the observations and recommendations of Bem, Palmer & Broughton (2001), Storm, Tressoldi & di Risio (2010), and Baptista, Derakhshani & Tressoldi (2015).

Studies will be included that meet the following criteria:

1. The study must be a ganzfeld ESP study (i.e. adopting a telepathy, clairvoyance or precognition method) conducted in a laboratory (i.e. not online).
2. **'Ganzfeld' is defined as follows:**
   a. The participant is seated or reclined comfortably in a quiet room.
   b. The participant is presented with a homogenous red visual field, for example by wearing translucent eye-shields, or being instructed to close their eyes while a red light is shone in their face, or wearing a VR headset. White noise or similar unpatterned auditory stimulation is played to the participant (e.g. through headphones) during the impression period.
   c. An impression period duration of between 20 minutes to 40 minutes is recommended.
   d. The target is randomly selected from a group of four possible orthogonal **visual** targets (e.g. postcards[2], art prints, film clips), such that the mean chance expectation of correctly identifying the target (a 'hit') is 25%. It is recommended that several groups of four orthogonal targets be assembled to form the target 'pool', so that the target set for each session is randomly selected from the larger pool.
   e. In the case of a clairvoyance or telepathy experiment, the selected target is displayed in a non-adjacent room to minimize the possibility of sensory leakage of target information.
   f. The participant's mental impressions ('mentation') during the impression period are recorded during or immediately after the impression period.
   g. Judging (i.e. rating the similarity between the mentation and the four possible targets) is done in a way that precludes cues from differential handling or viewing of the selected target. For clairvoyance and telepathy designs with paper targets, a duplicate set of targets should be used for judging. Judging may be done by the participant or by an independent judge. In either case, this judging must be done while the judge and anyone else in contact with the judge, including the experimenter, are blind to the actual target identity. No ties are permitted.
   h. The target identity is revealed to those involved in judging only after the results of the judging are recorded.

[2] Bem & Honorton's observation of higher hit rate with dynamic targets has not been confirmed in subsequent studies (Derakhshani, personal communication).

3. It is recommended that the participant undergoing ganzfeld stimulation is selected according to one or more of the following criteria:
   artistic, prior psi experience or training, practice mental discipline, belief in psi. It is also recommended that a demographic questionnaire is administered to record the presence or absence of these factors for each participant.
4. It is recommended that the participant undergoes a relaxation procedure prior to undergoing ganzfeld stimulation.
5. The study must be initially prospectively registered **after** the initial registration of this meta-analysis. The study must be registered on the KPU registry or on another study registry with comparable registration practices. Studies registered on other study registries will be included if (a) the registration information is adequate to evaluate these inclusion criteria, (b) the registration is unambiguously made irreversibly public before data collection begins, and (c) Prof. Watt becomes aware of the registration and is able to make the decision about including the study before the study results are known.
6. The final results for an included study results may be unpublished at the time of the final analyses for this meta-analysis. ('Published' is defined as reporting in conference proceedings, placing a description of hypotheses, methods and results on an online platform such as SSRN or f1000, or reporting the study in a scientific journal.) The minimum data that must be provided is the overall number of hits and trials in the study and the numbers of hits and trials for relevant subgroups identified in the registration for the study.
7. The study must be conducted at a university or established non-profit research institution, and have ethical approval where the institution has an ethical review process. Where no such process is supported, the study must conform to ethical guidelines such as those provided by the British Psychological Society, the American Psychological Association, or the Parapsychological Association.
8. The study must be directly conducted or closely supervised by an experienced researcher (defined as a researcher who has previously reported as first author peer reviewed hypothesis-testing research, e.g., online, at a conference, or in a scientific journal).
9. It is recommended that the study procedures include duplicate records handled in a way that makes it difficult for any one person (experimenter or participant) to intentionally or unintentionally alter the results without detection.
10. In the case of telepathy designs, the sender and receiver must surrender any mobile communication devices to prevent deliberate or inadvertent (e.g. sender pocket-calling) leakage of information about the target identity during the session.
11. The study must not have design flaws that could reasonably be expected to significantly compromise the results, or have an unusual design that would make the study not comparable to other studies.

List of Included Studies

The decision to include a study in the meta-analysis and the classification of the study as having artistic participants and as having a telepathy, clairvoyance or precognition design will be made and recorded at the time the study is initially registered and before any data are collected for the study. The list of studies to be included will be maintained on the KPU Registry (https://koestlerunit.wordpress.com/study-registry/registered-studies/) as a separate public document associated with this meta-analysis registration. The list will be updated as each new study is registered. The list of included studies may also have qualifications for the inclusion of data from a study. Any qualification will be prospectively specified at the time the study is added to the list. For example, for an experiment comparing a standard test condition with a novel test condition, it may be specified that the data from the novel test condition will be excluded from the primary confirmatory analysis. Or, if a study has artistic and non-artistic participants, the subset of artistic participants will be included in the analysis of artistic participants.

Protocol Deviations

If the registered number of trials is not obtained for one or more studies, the analyses will be conducted with two steps. The first step will be an analysis that includes all of the available data for all the studies that were planned to be included. If this analysis gives significant results, a second analysis will be done that applies the principle of handling protocol deviations conservatively (with the assumption that psi does not occur and that potential methodological biases from protocol deviations did occur). The second analysis will add data with specified scoring rates to obtain the total number of trials that was expected from the list of included studies. The added data for trials that were known to have not been conducted will have chance scoring. The total number of trials known to have not been conducted for all studies will be determined and trials added to obtain that total. The added trials will have scoring ≤ 25% and be as close to 25% as possible for the total number of these added trials. The added data for trials that were not provided for the meta-analysis but were known to have been conducted or may possibly have been conducted will be assumed to have had unfavorable results that are significantly below chance. The number of trials that were possibly conducted but not provided will be conservatively estimated for each study. For each such study, these added trials will have a scoring rate that produces psi missing with $p = .025$ one-tailed (or as close as possible) for the estimated number of trials possibly conducted but not provided for the study.

All other types of protocol deviations will be similarly conservatively handled for the second analysis. If the first analysis gives significant results but the conservative analysis for protocol deviations does not give significant results, the meta-analysis will be considered to have produced inconclusive results due to protocol deviations.

8. The power analysis or other justification for the number of participants and trials.

Different researchers have come up with different effect sizes in analyses of ganzfeld studies. The most recent peer-reviewed retrospective meta-analysis of ganzfeld ESP studies is Storm, Tressoldi and di Risio (2010), who meta-analysed 30 studies from 1997-2008. After excluding one statistical outlier (STDR's decision on grounds of homogeneity) and one study with 8 target options, the pooled number of hits was 483 out of 1498 trials, giving a 32.2% hit rate (binomial z = 6.44).

However, Baptista, Derakhshani & Tressoldi (2015) reported that the heterogeneity of the results could be explained by treating unselected subjects and selected subjects as two distinct groups. The hit rate for selected subjects was 40.1% (95% confidence interval 36.5% – 43.7%) and for unselected subjects 27.3% (95% confidence interval 24.3% – 30.3%). They noted that this hit rate for selected subjects was much higher than the hit rates of 31% to 34% found for selected subjects in earlier meta-analyses. If the hit rate of 27.3% for unselected subjects is accurate, it is not practically possible to conduct an adequately powered study with unselected subjects (2254 trials for .80 power). The distinction between selected and unselected subjects is difficult to address precisely because the subjects who volunteer for psi experiments are likely self-selected.

The strongest and most consistent effects in ganzfeld research have been with artistic participants (visual artists, musicians, actors, and writers) (Baptista, Derakhshani & Tressoldi, 2015; Dalton, 1997). This is a subset of selected participants. The average hit rate for this group is 41% for seven studies by different experimenters. The later studies were generally viewed as confirmations of the first study by Schlitz and Honorton (1992). This group, all of which used a telepathy design[3], is appropriately evaluated separately in a prospective meta-analysis.

If two tests are done to evaluate a hypothesis that is true, the probability that at least one of the tests will produce a significant result can be substantially higher than the power for each test alone. For example, if one test has a power of .50 and the other has a power of .60, the probability that at least one of the tests will be significant is .80 (which is one minus the probability that both tests will not be significant, or 1 - (.50 x .40)). This result is based on the assumption that the tests are independent. If the tests are correlated, then the power will be less than if the tests are independent. If the tests are perfectly correlated, there is no increase in power from doing two tests. For the planned ganzfeld meta-analysis, the two tests will be correlated, but the degree of correlation will depend on the number and size of the studies and cannot be reasonably estimated prospectively. For example, if each test has a power of .80, the probability that at least one test will be significant will be in the range of .80 to .96 (for a correlation of one to a correlation of zero). The actual power for the two tests for the ganzfeld meta-analysis will probably be somewhere in the middle of the range. The benefit from doing two tests is greatest when both tests have low power and is less when one or both tests

[3] One study (Morris, Dalton, Delanoy & Watt, 1995) had 32 clairvoyance trials (34% hit-rate) and 64 telepathy trials (31% hit-rate).

have high power. The benefit can also be substantial if one test is for an identified subgroup that has larger effects than the rest of the data.

For the confirmatory analysis of the total pooled data, given the variation in previous results, it was decided to design for a high power (.95) for a hit rate of 30% with alpha = .04. This hit rate is estimated to occur with a mixed sample of participants (selected and unselected). A 30.0% hit rate is approximately the lower 97% one-sided confidence interval for the 32.2% (483/1498) hit rate reported by Storm et al. (2010) (see Appendix). Setting alpha to .04 gives an appropriate adjustment for two analyses with a much stronger expected effect in the other analysis and will result in higher power for the outcome of both tests.

The sample sizes needed for various hit rates and power levels and the power for the total pooled data sample size (**921 trials**) are shown in table 1 below. If the studies have predominantly selected subjects and the previously reported hit rates are correct, the meta-analysis will have extremely high power. However, if the subjects are predominantly unselected, the power may be marginal to weak.

Table 1: Number of trials for different hit rates and power for the pooled total

Hit rate %	Number of trials for alpha = .04 one-tailed			Power for 921 trials
	Power=.80	Power=.90	Power=.95	Alpha=.04
27	3225	4431	5555	.36
28	1457	1990	2505	.62
29	826	1135	1419	.84
30	**535**	**735**	**921**	**.95**
32	282	382	482	.998

For the confirmatory analysis of the artistic subset, the number of trials for different hit rates and power are given in table 2 below. A hit rate of 37.0% is approximately the lower 95.5% one-sided confidence interval for the 41.4% (152/367) hit rate reported by Derakhshani (2012) for 7 ganzfeld studies with artistic participants (see Appendix). Therefore the target sample size for the artistic group is **234 trials**, corresponding to .95 power to detect a 37% hit-rate. A hit rate substantially higher than 37.0% is likely if the effects in the previous research are valid. Also, tests with lower power contribute to higher power for the outcome of two tests.

9. The methods for randomization in the experiment.

The methods for randomization in each experiment will be detailed in the study registration document for the experiment.

Table 2: Number of trials for different hit rates and power for the artistic group

Hit rate %	Number of trials for alpha = .01 one-tailed			Power for 234 trials
	Power=.80	Power=.90	Power=.95	Alpha = .01
30	802	1038	1261	.27
31	561	724	884	.39
32	418	539	657	.52
33	320	418	509	.64
34	256	331	400	.76
35	210	277	331	.84
36	175	230	277	.91
37	**150**	**195**	**234**	**.95**
38	127	168	203	.975
39	110	144	175	.988
40	97	127	151	.995
41	**87**	**110**	**134**	**.998**
42	77	97	120	.999
43	70	87	107	1.00

Note: The power calculations were from the G*Power program (Faul, Erdfelder, Lang, & Buchner, 2007) available at http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3.

10. A detailed description of the experimental procedure.

The procedure for each experiment will be detailed in the study registration document for the experiment.

11. Publication policy. To avoid divergent expectations about publication and authorship, the Cochrane Handbook (Ghersi, Berlin & Askie, 2011) recommends that the prospective meta-analysis protocol includes a statement of publication policy. This prospective meta-analysis does not introduce any limitations on the experimenters for the included studies reporting and publishing their studies. The authors for publication of the prospective meta-analysis will be those who substantially contributed to the planning, data analysis, and writing of the meta-analysis. The experimenters for the included studies will typically not be co-authors for this registration-based prospective meta-analysis.

References

Baptista, J., Derakhshani, M. & Tressoldi, P. (2015). Explicit anomalous cognition: A review of the best evidence in ganzfeld, forced-choice, remote viewing and dream studies. In Etzel Cardeña, John Palmer, and David Marcusson-Clavertz (Eds.) *Parapsychology: A Handbook for the 21st Century*. Jefferson NC: McFarland.

Bem, D. J., Palmer, J., & Broughton, R. S. (2001). Updating the ganzfeld database: A victim of its own success? *Journal of Parapsychology, 65,* 207–218.

Cooper, H., & Hedges, L. V. (2009). Potential and limitations. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed.) (pp. 561-572). New York: Sage.

Dalton, K. (1997). Is there a formula for success in the ganzfeld? Observations on predictors of psi-ganzfeld performance. *European Journal of Parapsychology*, 13, 71-82.

Derakhshani, M. (2012). On the statistical replicability of ganzfeld studies. Presented at the 55th Annual Parapsychological Association Convention, 2012.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191.

Ghersi, D., Berlin, J., & Askie, L. (2011). Chapter 19: Prospective meta-analysis. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from http://www.cochrane-handbook.org.

Higgins, J. P. T., Deeks, J. J., & Altman, D. G. (2011). Chapter 16: Special topics in statistics. In: Higgins, J. P. T. & Green, S. (editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from http://www.cochrane-handbook.org.

Kennedy, J.E. (2013). Can parapsychology move beyond the controversies of retrospective meta-analysis. *Journal of Parapsychology*, 77, 21-35. Available from http://jeksite.org/psi/jp13a.pdf.

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

Schlitz, M. J., & Honorton, C. (1992). Ganzfeld psi performance within an artistically gifted population. *Journal of the American Society for Psychical Research, 86,* 83-98.

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

Appendix: Calculation of binomial confidence intervals, by Jim Kennedy.

Binomial confidence intervals are from http://statpages.org/confint.html The Clopper-Pearson method used in this calculator is sometimes claimed to be overly conservative (wide intervals) but it apparently is the only method that always has at least the specified confidence interval and maintains a nested relationship among different intervals.

For overall Storm et al. 2010, 483/1498 = 32.2%, 95% CI 29.88%-34.7%

.3068 is lower 90% 1-sided CI
.3025 is lower 95% 1-sided CI
.3005 is lower 96.5% 1-sided CI
.2997 is lower 97% 1-sided CI

For artistic Derakhshani 2012, 152/367 = 41.4%, 95% CI 36.3%-46.6%

.3803 is lower 90% 1-sided CI
.3711 is lower 95% 1-sided CI
.3699 is lower 95.5% 1-sided CI