A Forced-choice Precognition Experiment with Selected Cohorts

Peter A. Bancel, Jocelyne Boban, Anaïs Bensahra and Mario Varvoglis

Institut Métapsychique International, Paris

Abstract

Objective. We report a pre-registered forced-choice precognition study using a protocol designed to enhance psi effects. Two selected participant populations were compared: a cohort of experienced meditators and a cohort self-selected by the mere interest to participate in the study.

Method. An Internet-based experimental platform was developed to allow participants to complete sessions at home using software installed on their personal computers. We call the platform Psi@Home. Participants completed multiple sessions of 20 forced-choice trials. 80 sessions per cohort were collected for the formal study and hypotheses for each cohort were: 1) increases of the variance of session hit rates for each cohort; 2) a significantly higher variance for the meditator cohort; 3) increases of the total hit rates for each cohort. Hypotheses (1) were confirmatory, and the others were exploratory.

Results. The variance difference hypothesis was confirmed (p=.04), and the other hypotheses did not surpass p=.05 one-tailed significance. 90 sessions of preliminary data, whose collection was stipulated by pre-registration, showed strong increases of session variance (p=.00004).

Conclusion. The registered hypothesis tests did not produce evidence for a psi effect in the formal registered study. However, evidence for an increased variance of session hit rates did appear strongly in preliminary data that was specified in pre-registration and collected using the same protocol. Differences in participant attitudes during the two periods of data collection may account for the discrepancy.

Keywords: *psi, precognition, forced-choice, selected subjects, meditation, psi-missing*

Highlights:

- A novel platform for running experiments with selected cohorts was developed.
- The Internet-based Psi@Home platform is designed to be available to outside researchers.
- The platform completed a study of 160 sessions with 47 participants within a month.
- A strong psi effect was found in preliminary data, but not in the formal study, possibly due to uncontrolled psychological variables.
- The strong psi effect indicates a mixture of psi-hitting and psi-missing across sessions.

Introduction

In recent years, new meta-analytic studies have strengthened the evidence for psi effects from free-response protocols such as remote-viewing and the ganzfeld (Tressoldi & Storm, 2024; Tressoldi & Katz, 2022; Storm & Tressoldi, 2020). The evidence is further supported by modeling studies that control for publication bias and other methodological issues detrimental to meta-analysis (Bierman et al., 2016; Bancel, 2018). For studies that use other protocols, such as forced-choice in which subjects register a choice among a predefined set of randomized alternatives (for example, guessing the outcome of a coin flip), the cumulative results are quite positive, given the very different procedures and effect sizes that these protocols entail (Storm & Tressoldi, 2023; Bem et al., 2015).

Despite the accumulated evidence, the challenge of replicability in parapsychology remains. The success of one-off experiments is far from guaranteed, even when studies are presumably well-powered. For example, some recent attempts at registered, large-scale replications have failed to produce an effect (Schlitz et al., 2021; Kekecs et al., 2023). This state of affairs is not new and the coexistence of strong evidence and replication uncertainty has been recognized in the psi

literature for a long time, particularly for forced-choice and micro-PK protocols (Bem et al., 2015; Bosch et al., 2006). For those who do accept the evidence, the situation highlights the difficulty in creating the necessary conditions for psi to occur, or – in some interpretations – leads to the idea that psi is real, but somehow resistant to replication (Walach et al., 2022). For many in mainstream science who are skeptical of the psi hypothesis, assurance of replication is a sine qua non for accepting the reality of an effect and, for these researchers, the replication difficulties derail any consideration that the data anomalies represent real phenomena (Rouder & Morey, 2011).

Over the years, there has been much effort to find better methods to produce psi effects in the laboratory (Palmer, 2015). Among the successful efforts are the afore-mentioned ganzfeld and remote viewing protocols. These rely on techniques to induce favorable psychological states and are among the most effective methods used to produce psi. However, notable drawbacks include the high cost in human resources and the considerable tacit knowledge required of experimenters. Even moderately well-powered ganzfeld experiments are quite onerous, so that any progress beyond adding to the evidence tends to be incremental, at best. Consequently, single one-off replications of high power are extremely resource intensive and are rarely attempted (for the report of a recent large study see Watt, 2024).

Other protocols, such as forced-choice GESP, micro-PK (typically with random sources such as hardware RNGs) and physiological presentiment have higher data rates and are often less time-consuming (Jahn et al., 2007; Radin & Pierce, 2015). They are also able to address a wider range of research questions. But the results are more volatile and success often relies on the efforts of skilled experimenters (Schlitz et al., 2006; Varvoglis & Bancel, 2015). This seems to preclude a recipe for general replication and even confounds the interpretation of data because it begs the

question of whether psi is sourced in the subjects, in those running the experiments, or a synchrony of the two.

In summary, nearly a century of psi research has yielded a variety of methods which have produced plenty of evidence. And yet the methods are unsatisfactory because they are either too resource intensive and ill-suited to process-oriented work, or they produce effects that are difficult to replicate. In consequence, trade-offs between effect size and data rate; reliability and design flexibility; and cost and replicability impede progress.

The experiment reported here is part of a long-term effort at the Institut Métapsychique International (IMI) to address these problems. The program focuses on developing induction techniques and data-collection methods that are faster and easier to implement, while allowing for flexible experimental designs. The strategy is to bring together the most fruitful elements of diverse psi protocols in order to mitigate the negative trade-offs and allow for experiments that are flexible, reliable and more practical in terms of resources.

Our preferred framework for this program is the forced-choice approach because it allows for higher per session data rates and affords rich data structures for subsequent analyses. Of course, as we know, the potential disadvantage of this approach is the risk for much smaller and more volatile effect sizes (Storm & Tressoldi, 2023). Particularly if one contrasts the ‘subject optimization’ procedures of remote-viewing or ganzfeld trials, with the repetitive task-feedback cycle of forced-choice protocols, it seems plausible that the latter can induce psi-inhibitory conditions: boredom, loss of motivation, stress about trial outcomes, and so forth.

Yet, while plausible, this understanding of the low effect sizes in forced-choice experiments lacks unequivocal empirical support. To truly assess its validity, we need protocols that can efficiently collect large amounts of data while systematically modulating appropriate psychological

variables. A key objective of our research program is to study this question by providing researchers with a flexible yet powerful tool to test hypotheses concerning psi correlates. At the same time, our intuition is that putative psychological inhibitions to psi performance can be remedied by appropriate subject-selection, and by designing tasks that are embedded in an engaging and immersive environment that favors psi-conducive states.

In short, then, the overall aim of our program is to explore how to integrate psi-conducive factors into data-efficient forced-choice protocols. The practical objective here is to develop reliable protocols that would not only speed progress, but also render psi research more accessible to outside researchers. A more conceptual objective is to attempt to resolve the tension between views that consider psi's elusiveness to be merely circumstantial versus those that treat it as fundamental. In particular, a current proposal considers psi effects to be inherently elusive and by their nature resistant to replication (von Lucadou et al., 2007). In this latter view, attempts to develop reliable protocols are likely to fail (Walach et al., 2022). Our working assumption is that this view is incorrect, or at least too categorical and our hope is that the protocols we are developing will provide clarification on this issue.

Two factors that we focus on are experientially immersive psi tasks and the selection of subjects. While these have been studied previously, we make some innovations and employ a design that attempts to optimize both in a forced-choice protocol. The immersive presentation we use is based on prior development work in our laboratory (Varvoglis, 2014; Varvoglis, 2019). It has been adopted for the current experiment and is described in the Methods section. Selecting subjects for their potential to produce psi effects has a long history in parapsychology. Instances of gifted subjects who have performed well under a variety of circumstances are well-documented (see for example May & Marwaha, 2018). However, an obstacle to replication with

gifted subjects is that they are rare, and often unable (or unwilling) to produce effects on demand. Selecting persons by traits that are assumed or hypothesized to favor psi performance is another avenue that has been studied. While there are indications that selection by reported or measured traits may enhance results, to our knowledge no inventory or survey reliably predicts psi performance.

For the present experiment we selected two subject cohorts based on life experiences and attitudes thought to be associated with psi performance. The two criteria are extensive experience in meditation, and the mere interest in psi research. The practice of meditation has long been associated with psychic abilities, dating back to at least the writings of Patanjali (Woods, 1927). Parapsychological studies with meditators have given indications of enhanced performance, but there is as yet no conclusive evidence that meditators outperform the general population (Roney-Dougal, 2015). One difficulty that arises in parapsychological studies with meditators is assessing the degree and depth of peoples' meditation practice. An approach to this problem is presented in the Methods. The second cohort we study is a group of persons selected for having an interest or openness to psi phenomena. In parapsychology, 'sheep' is an informal term for people who profess favorable attitudes towards psi, ranging from mere interest to openness to belief (as opposed to 'goats', or those firmly skeptical or disbelieving that psi effects are real). The so-called sheep-goat effect hypothesizes that sheep will perform better than goats on psi tasks. Evidence for the sheep-goat effect is encountered widely in the literature and the idea is generally accepted in the field. However, meta-analyses do not provide conclusive support for the effect and better data is needed to resolve the question (Storm & Tressoldi, 2017; Lawrence, 1993). A subsidiary goal of this experiment is to contribute data to test the sheep-goat hypothesis against future data from a skeptical cohort.

Finally, the study takes into consideration that misdirection of psi effects can contribute to the variability of results. In forced-choice experiments, true psi effects may produce data that deviates opposite to the intended target direction. There is considerable evidence for "psi-missing", as it is called (Rhine, 1969; Storm & Ertel, 2001; Carpenter, 2004), and its presence can weaken the statistical power of directional tests. Therefore our psi hypotheses include tests that have been devised to optimize statistical power under models of psi-missing.

Method

We developed and tested a platform for running home-based studies with selected cohorts that allows a modular approach to experimental design. By modular we mean that the platform allows for the independent design of three essential experimental elements– a cohort; the researchers; and an experimental task. In this section we describe 1) the configuration and technical aspects of the platform; 2) a computer application used by cohort members to run at-home sessions; 3) the process of cohort selection; 4) the experimental hypotheses and pre-registered data analysis; 5) the procedure for running studies with cohorts.

The Psi@Home platform.

The experimental platform, which we name Psi@Home, consists of a downloadable application, its connection to a web-based server, and a web-site used for cohort recruitment and management. The application, described below, is bundled into a custom installer package for distribution to cohort members. The package includes custom software to manage login, security features, and data communication with the server, as well as the main application for running experimental sessions. Data records are uploaded in real-time to the cloud server at Amazon Web Services (AWS). The AWS account serves as a repository for all experimental and cohort login data, and provides storage and download links for different versions of the application bundle.

This provides a way to deploy and maintain multiple experiments from one integrated platform. The AWS is configured to manage multiple accounts for an unlimited number of experimenters and to run studies in parallel.

The website, https://imiresearch.fr , is the public face of the platform. It is used to publicize experiments, recruit participants and manage cohorts. The website provides general information about the research, sign-up forms for recruitment and information for individual cohorts or current experiments. The website is housed on the commercial platform Wix, and allows for cohort management via email and form capabilities linked to Google Workspaces. Administrative accounts allow researchers to collaborate with specific cohorts and schedule online video meetings with cohort members. A dedicated Zoom account is owned by the platform for this purpose.

The Selfield application.

At the heart of Psi@Home is an application that cohorts use to run at-home experimental sessions. A key feature of the platform is the ability to create and deploy different applications. This permits a wide flexibility in designing studies adapted to particular research questions. In this work, we use a binary test of precognition that has been designed and tested at the IMI. Named the Selfield for its immersive quality, the test fluidly presents successive forced-choice trials via an engaging and relaxing graphical interface. The Selfield is designed to enhance participants' attention to the task and lessen boredom and loss of motivation. Laboratory tests have shown that participants' experience with the Selfield is almost uniformly positive (Varvoglis et al., 2019). Similarly high ratings for pleasantness were found with the Psi@Home cohorts, as discussed in the Results section.

Each formal Selfield session consists of 20 trials where participants interact with a graphical "target container" which is presented on the computer screen as a luminous, floating blue sphere.

Participants are asked to choose the moment to reveal a target hidden inside the container. At a moment of the participant's choosing via a keystroke, the container is revealed to be either empty, or instead, to contain the striking image of a personage. An empty container is considered a 'miss' and finding a personage, a 'hit'. Each instance of hit or miss is determined by a pseudorandom process that is seeded anew for each trial using input from the milli-second timing of two participant keystrokes (one which readies the choice and the second that executes it). After the reveal, the Selfield proceeds to the next trial until all 20 trials are completed. A session lasts about 15 minutes, but participants take as long as they wish to complete the trials.

Participants were instructed that the experiment tests for psi and that their intention should be to "meet as many personages as possible", that is, to obtain hits. The null expectation in the absence of psi is a 50% hit rate, which is the null expectation of the random generator. The psi effect usually tested by the Selfield is for an amplification of the hit rate. The task is considered precognitive because the pseudo-random process to determine a hit or miss occurs only after the user input (the keystrokes). A detailed description of the application is found in Varvoglis, 2019.

Cohort recruitment.

Cohort recruitment proceeds in two steps. People are contacted by various means and directed to the website where they can sign up as cohort candidates. To join a cohort, candidates must complete a try-out of Psi@Home. After completing two try-out sessions, those who wish may formally join the cohort. The current study established three cohorts: experienced meditators, a general public "Open" cohort, and a third Psychic Arts cohort which was combined with the Meditator cohort for the formal experiment.

The recruitment procedure plays a central role in the conception of Psi@Home. The goal is to motivate participants and create a connection to the project, and to cull unmotivated candidates before experiments are run. Candidates install the Psi@Home application with a team member

during a video meeting. The meetings last about 45 minutes and are intended to enhance the candidate's trust, motivation and familiarity with the project. After testing the software together, the team member explains that two try-out sessions need to be completed within a week's time. These sessions are of the same length and format as the sessions that participants will complete during formal studies. When completed, candidates receive an email to welcome them to the cohort and inform them that they will be invited to participate in future experiments. The recruitment procedure builds participant motivation to contribute to the project's studies. We believe that the quality of these interactions can play a role in experimental outcomes. A long-term goal is to build a database that can help assess how qualitative procedures such as recruitment may impact the results of psi experiments.

The procedure was first tested with 5 meditators personally known to the PI (principal investigator – PAB). Their feedback allowed to refine and clarify the process from a user perspective. Data from the resulting 10 try-out sessions were used to test analysis procedures in preparation for the formal experiments to come. Members of the meditator cohort were selected from a community of buddhist practitioners that maintains a database of individuals' progress. All had 15 to 40 years experience, maintained daily home practice, practiced the same techniques of mindfulness, visualization and mantra, had completed many group retreats, and most were meditation teachers. The PI has similar experience and knew personally most of the cohort. The cohort's depth and homogeneity of meditation experience, as well as the familiarity shared by the PI, is a rather unique instance in psi studies with meditators.

The Psychic Arts cohort consists of persons involved professionally in mediumship or clairvoyance practices, or persons actively involved in training for these or similar psychic arts. Many were recruited among members of the International Remote Viewing Association (IRVA)

via presentations by the PI or emailings to IRVA members. The Open cohort for the general public was solicited from email lists of the IMI's sister association (Friends of the IMI) which is active in educational outreach about psi in France, and from online presentations in English by the PI.

Recruitment for the Meditator cohort was conducted July to early October 2022. Of 81 persons contacted, 20 installed the Selfield application and 19 joined the Meditator cohort. The Meditator cohort generated 38 try-out sessions during recruitment.

Recruitment for the Psi Arts cohort was conducted August to October, 2022. A total of 14 persons installed the Selfield application and 11 joined the Psi Arts cohort. The Psi Arts cohort generated 20 sessions during recruitment.

Recruitment for the Open cohort was conducted August to October, 2022. 27 persons installed the Selfield application and 23 joined the Open cohort. The Open cohort generated 58 sessions during recruitment.

We estimated that a cohort pool of about 25 members would be needed to complete each study. Because the Meditator and Psi Arts cohorts were below this mark at the end of the 3-month recruitment period, a decision was made to combine the two cohorts for this first experiment (in the following, the combined cohort is referred to as the Meditator cohort, unless otherwise stated). Of the 30 members of the (combined) Meditator cohort, 24 joined the experimental study (22 female and 2 male), and all 23 Open cohort members participated (18 female and 5 male).

Hypotheses and analyses.

Prior to the experiment, data from try-out sessions was analyzed in order to finalize hypotheses. Three types of 1-tailed hypothesis tests were set:

- The variance of session hit counts for Meditator and Open cohorts will exceed null expectation (with $p<.05$). Confirmatory; noted in the pre-registration document as H1 and H2.

- The Meditator cohort variance will be greater than the Open variance at the 0.05 level. Exploratory; pre-registered as H5.
- Cohort hit rates will exceed >50%, at the 0.05 level on a direct binomial test. Exploratory; pre-registered as H3 and H4.

The experiments were pre-registered with the Koestler Parapsychology Unit Study Registry (http://www.koestler-parapsychology.psy.ed.ac.uk/Documents/KPU_Registry_1072.pdf).

The number of sessions for each cohort experiment was set to 80, and participants were asked to complete 4 to 6 sessions of 20-trials each. We allowed for the collection of more than 80 sessions per cohort, but the formal hypothesis tests were performed on the first 80 cohort sessions only, as per our pre-registered procedure.

The variance hypothesis tests return a p-value for the zero-mean variance of session hits. This is defined as the sum of squared session z-scores, where z^2=4/N*(hits-N/2)^2, and N is the number of trials (see the Appendix for precise definitions of these terms). Nominally N=20, but occasionally sessions recorded fewer trials due to participants' intermittent WiFi connections (about 10% of sessions). To maintain nearly equal weights of sessions, the registered protocol stipulated that sessions with less than 17 recorded trials should be discarded (about 5% of sessions). The sum of z^2 over all sessions approximately follows a chi-squared distribution with NS degrees of freedom (NS being the number of sessions). P-values can thus be estimated analytically from the appropriate chi-squared distribution function and we use these as checks on more precise Monte Carlo (MC) estimates of the p-values. Full details are available in Appendix and in the protocol pre-registration document (http://www.koestler-parapsychology.psy.ed.ac.uk/Documents/KPU_Registry_1072.pdf).

Study procedure.

Studies begin with an email invitation to cohort members that describes the study launch date and duration, the requirements for participation (e.g., the number of sessions to complete), and gives a link to an online consent form. Cohort members join the study by accepting the invitation, completing the consent and self-installing a software update of the Selfield application. On the launch date, participants are invited to an optional collective online video gathering to clarify any remaining questions and provide a final encouragement to the group. The updated Selfield applications are then activated and participants are free to contribute sessions at the times of their choosing. Our design is for studies to last 4 to 6 weeks, with participants individually contributing 4 to 6 sessions. Participants receive an email reminder if they lag in completing sessions, but care is taken not to pressure people for results. Studies assume that some will complete less than the requested number of sessions and allowance for this eventuality is incorporated into the study design. At the study's end, participants are invited to an optional closing video call where they can be thanked and express their experiences to the group. As a final step, participants fill out a brief online feedback survey to assess their experiences.

Results

A major objective of our study was to test the functionality of the Psi@Home platform and assess its potential for carrying out psi studies quickly and efficiently. The studies ran smoothly, without major difficulties or unexpected problems, and the demands on the experimental team were less than we anticipated. Study invitations were emailed on October 10, 2022 and we were able to launch a week later. Data acquisition for both cohorts, comprising 80 formal sessions each, was completed within 30 days. The Psi@Home platform surpassed expectations for study execution and management.

A second objective was to assess the recruitment procedure for establishing cohorts. Through early October 2022, we received 104 website submissions for cohort candidacy, of which 61

(59%) installed the Selfield application during an online video call. Of those who installed the application, 53 completed the try-out sessions and joined a cohort (87%). Altogether, 47 cohort members participated in our formal study (89% of the cohort). Roughly speaking, we converted about half of contacts to cohort membership, and nearly 90% of the cohorts were available for the experiment.

Pre-registered confirmatory hypotheses

The confirmatory hypothesis of an increase in session variance was not confirmed for either cohort. The session zero-mean chi-square for the Meditator cohort was slightly greater than mean expectation ($\boldsymbol{X}2(80) = 89.43$; p = .22; 40k Monte Carlo iterations). The session zero-mean chi-square for the Open cohort was moderately lower than mean expectation ($\boldsymbol{X}2(80) = 59.43$; p = .96; 40k MC iterations).

Pre-registered exploratory hypotheses

Hypothesis 2 (larger session variance for the Meditator cohort) was confirmed ($\Delta\boldsymbol{X}2(80) = 30.04$; p = .04; 100k MC iterations).

The hypotheses 3, a positive bias of the hit rates, was not confirmed for either cohort. The Meditator study generated 780 Hits on 1597 trials (Hit Rate = 48.8%; exact binomial p = .83, one-tailed). The Open cohort study generated 825 Hits on 1593 trials (Hit Rate = 51.8%; exact binomial p = .08, one-tailed).

Analyses of preliminary data

The planned pilot study examined the first 10 try-out sessions from the Meditator cohort. The goal of the pilot study was to test cohort management procedures and the analysis algorithms for hypotheses 1 and 3 (session variance and total hit rate). However, the data also gave indications of a psi effect. The session zero-mean chi-square for the pilot sessions was ($\boldsymbol{X}2(10) = 18.2$; p =

.016; 40k Monte Carlo iterations). For these sessions, there were 89 hits on 200 trials (hit rate = 44.5%; exact binomial p = .948, one-tailed).

The pilot result prompted an interim analysis of recruitment data when 50 sessions (contributed by both cohorts) had been accumulated. The session zero-mean chi-square for the 50 sessions was highly significant ($\boldsymbol{X}2(50) = 94.83$; p = .000085; 2M MC iterations). There were 506 hits on 994 trials (hit rate = 50.1%; exact binomial p = .318, one-tailed). This analysis was subsequently used as the basis for the pre-registered protocol, which set the number of sessions for each cohort to 80, and designated the variance test as a confirmatory hypothesis.

The analysis was updated for all 90 preliminary sessions that were completed by cohort members who participated in the study ($\boldsymbol{X}2(90) = 150.72$; p = .00004; 2M MC iterations). There were 905 hits on 1789 trials (hit rate = 50.6%; exact binomial p = .295, one-tailed).

Discussion

Our report addresses two research objectives. First, we tested the functionality of a new platform, Psi@Home, whose broad purpose is to facilitate the design and execution of psi experiments. Second, we used the platform to run a study with the goal of eliciting evidence for a psi effect and comparing two cohorts. We discuss the outcomes of each of these objectives in turn.

Assessment of the Psi@Home platform

In terms of functionality, the Psi@Home platform met all of our design goals. The Selfield application was successfully integrated to our cloud-based data management system and the installer package we designed allowed for easy installation by the individual cohort users. The website created for cohort recruitment and management worked well for scheduling and email communication throughout the recruitment process, both within the project team and between

team and cohort members. We have processed over a hundred contacts and guided scores of people through the process of joining cohorts. The recruitment process did prove to be somewhat longer and more time-consuming than hoped. Outreach did not generate contacts at the rate we hoped and video calls required more effort than expected. However, the process was manageable and the value of personal contact between team members and participants compensated the effort. In particular, the interactions with cohort members successfully clarified and motivated their participation. This was reflected in the results of a feedback survey. A question :“Were the instructions and description clear enough ?”, resulted in an average score of 4.8 on a scale of 1 to 5 with 5 being “Very clear”. Our goal of a positive user experience was also met. A survey question “Did you enjoy using the Selfield app?” yielded an average response of 4.4; and “Would you recommend this to others?” yielded 4.5.

The decision to have candidates run two full experimental sessions before committing to a cohort was a valuable feature of the recruitment process. It allowed participants to have a good sense of how to operate the application before participating in a study, and served on a few occasions to cull candidates whose motivation was short-lived. It also allowed a thorough verification of the technical integrity of each installation, which was important given the variety of computer configurations we encountered. Clearly, recruitment was limited by the restriction to the Apple OS platform, but this can be rectified by porting the application to a PC compatible format in the future. Indeed, of the initial contacts who did not do the installation, most were willing but lacked access to an Apple computer.

The clearest measure of the platform’s success was the ease and rapidity of running the two formal studies. Three steps were required of each cohort member: response to an email invitation to join the study; the submission of a consent form and completing a software update; and the

accumulation of 4-6 experimental sessions. The steps were accomplished smoothly and quickly. After the invitations were emailed, the studies were ready to launch within a week. Once participants were informed of the launch, they began running sessions whenever they wished. The target of 80 sessions per cohort was reached in less than 30 days. During this time, the team monitored progress and sent a few reminders by email. There was very little further effort required by the experimental team and we attribute this success to the motivation and familiarity with the platform acquired by the cohorts during the recruitment process. This was precisely the outcome the project aimed for: to establish a pool of selected participants, experienced with the platform, who would respond enthusiastically to a subsequent call for study participation. Ultimately, the Psi@Home project is intended as a "user facility" for psi experimentation that is available to external research teams. This intent recalls the modular conception of the project whereby the research team is considered as one of three fundamental elements that compose a study. To this end, we have worked to make the platform user-friendly, so that experienced researchers can utilize it without too steep a learning curve, or the need for special technical knowledge. To test this aspect, the current study split the tasks of cohort and study management between cohorts, so that the PI (who designed all aspects of the platform) and two assistant team members without technical knowledge of the platform, worked separately with the Meditator and Open cohorts, respectively. We found that the Open cohort managers were able to use the platform efficiently after a brief introduction and training period. They managed cohort recruitment, and the formal study, with only occasional assistance from the PI, and the progress tracked that of the Meditator cohort. We conclude that the platform will be transferable to external researchers, either for running experiments with existing cohorts, or for establishing new

cohorts that can be used to study hypotheses of interest. The Psi@Home project will continue development work in this regard.

Formal pre-registered hypothesis tests

Turning to the results of data analysis, our two confirmatory pre-registered hypothesis tests did not reject the Null hypothesis. For each cohort, the defined session variance produced non-significant p-values (significance level of .05).

For exploratory pre-registered tests, the one-tailed difference of variances for the Meditator and Open cohorts was significant with p=.04. However, we are cautious about over-interpreting the result, given its closeness to the significance boundary and the fact that the corresponding confirmatory tests were non-significant. The exploratory tests of hit rates were non-significant. A noticeable difference between the Open and Meditator cohort hit rates (51.8% vs. 48.8%) yields a difference z-score of z=1.66, and a two-tailed p-value of p=.096 (see Appendix). We do not consider this post-hoc observation to be suggestive of an effect. Clearly, this first registered study failed to find support for the hypothesized effects of psi-hitting, or a mixture of hitting and missing, notwithstanding the marginal support for a variance difference of the cohorts. The apparent difference in variance between cohorts – marginally significant and in the hypothesized direction – may well be a false positive: the probability of one of the five tests returning a p-value of p=.04 or less is about 18.6%. It may be that the protocol did not evoke a psi effect at all; or, that an effect was too weak given the study size; or, that a psi effect obtained in a way not sensitive to our tests. More studies are needed to resolve these possibilities. However, the study results contrast with the data gathered during recruitment. Those data, collected with the same procedures and software, produced strong variance increases. We next discuss those results.

Assessment of the preliminary data

For the 90 recruitment sessions we find (Hit rate=50.59%, p=.32, trialN=1782; $\boldsymbol{X}2(90)=150.72$, p=.00004). The large variance is nearly 4 standard deviations from chance expectation (p=.00004 corresponds to a z-score of 3.94), and this is too extreme to ignore. One explanation consistent with the psi hypothesis is that a mixture of psi-hitting and psi-missing significantly increased the session variance but not the overall hit rate, and that psychological factors account for the lack of this effect in the formal study. However, before discussing this possibility, we address alternate explanations based on technical and methodological deficiencies.

It is conceivable that a deficiency of the Psi@Home software generates sessions with high variance that this explains the anomalous recruitment data. The null results of the registered formal study argue against this possibility. Two other data sets of comparable size that were collected during the study period also showed no variance anomaly. One was 74 extra cohort sessions collected after the registered N of 80 sessions per cohort was reached. A second was 65 sessions collected by a researcher who tested the Psi@Home platform independently during the recruitment period (with participants not from Psi@Home cohorts). Tests of variance for the data sets give insignificant p-values (respectively, $\boldsymbol{X}2(74) = 80.7$; p = .27; $\boldsymbol{X}2(65) = 60.0$; p = .66). The null results for nearly 300 sessions (the 160 registered sessions of the formal study; the 74 extra sessions of the formal study; the 65 independent researcher sessions) constitute a de facto control database, generated concurrently and under real-world conditions, that counters an explanation of a persistent software or platform malfunction. The Selfield application has had extensive prior use (Varvoglis, 2019) and such anomalies are not seen in earlier data. Further, contributions to the variance anomaly are distributed across many sessions and users (see below and Figure 1), so any malfunction would have to occur in multiple installations in the same manner. A few intermittent malfunctions cannot explain the variance. The variance anomaly is

not associated with the 5 sessions that dropped a few trials. Removing those from the 90 recruitment sessions doesn't impact the test ($X2(85) = 147.8$; $p = .00002$). Finally, we note that, although the software was updated just before the registered study, the update only changed a text file with a study identifier and did not alter the Selfield application itself. Identical software was used in both periods of data accumulation. These considerations lead us to conclude that the anomalously high variance in the preliminary data is not due to technical problems.

Methodological problems might also account for the anomaly. The formal analysis was peer-reviewed, registered and followed exactly. The express purpose of registration is to avoid questionable practices that can inflate p-values. The assurances of pre-registration is lacking for the analysis of recruitment data. However, analysis procedures were kept identical for both data sets and the Monte Carlo p-value was corroborated by exact calculations of p-values from theoretical chi-square distributions. That being said, an issue that does affect the stated p-value is multiple testing. We report the variance p-value of 90 recruitment sessions generated by study participants, but as mentioned above, multiple data sets were tested using both variance and hit rate statistics. In total 7 data sets were tested: recruitment data at 50 and then 90 sessions; formal data for 2 cohorts; the combined data for 2 cohorts; 74 extra cohort sessions; 65 sessions of an independent researcher. A Bonferroni adjustment multiplies the variance p-value by a factor 14 (accounting for the tests of hit rate and variance) which yields an adjusted value $p=.00064$. The adjusted p-value is still highly significant. The Bonferroni method is conservative, especially in this case where some data sets are not independent. We conclude that the variance anomaly is not an artifact of multiple testing or improper analytical procedures. We next consider the preliminary data under the psi hypothesis.

A model of psi-missing

As mentioned, psi missing refers a psi effect where outcomes deviate opposite to the intended target direction. Psi-missing has been discussed in the literature since at least the 1960s (Rhine, 1969; Rao, 1965). Recurrent psi missing can weaken statistical evidence for directional hypotheses such as tests of global hit rates, and render directional effects practically undetectable if hitting and missing are roughly balanced across sessions. The variance, however, may increase for a mixture of hitting and missing, and the zero-mean variance test is designed to detect this. It has been proposed previously (Timm, 1983) and used in other psi contexts (Storm & Ertel, 2001). Furthermore, studies in our laboratory (Varvoglis, 2013; Varvoglis, 2019), have found weak evidence for an elevated variance in forced-choice sessions. Therefore, there are both theoretical and empirical precedents for using the variance to test for mixtures of psi hitting and missing. A feature of the zero-mean test is its insensitivity to the proportion of hitting and missing: it returns the same result regardless of the relative frequency across sessions. We leave a full discussion of its properties to a future publication.

For the preliminary data, we saw that the variance test strongly rejects the Null hypothesis, even when adjusting for multiple testing. We note that a handful of sessions can account for this. Trimming the 6 most extreme sessions yields ($X2(84) = 106.8$, p = .044; Hit Rate = 49.8%, p=.50), which is a nearly complete attenuation of the variance anomaly. Trimming a seventh session increases the p-value to p=.08, which is above the conventional significance level of .05. Of the 6(7) trimmed sessions, 2(3) are in the missing direction. This observation supports the notion that the variance arises from a mixture of hitting and missing sessions. An outlier model that assumes a handful of extreme sessions – due to strong, but mixed, psi efforts of a few participants – can therefore explain the measured hit rate and variance. But it leaves unexamined whether the presumed psi effect appears only for those few sessions (less than 10% of the total),

or instead occurs broadly across the ensemble of participants and sessions. The question is an important one because the goal of the Psi@Home platform is to elicit effects broadly. We next address this issue by studying a competing model that assumes an effect for all sessions.

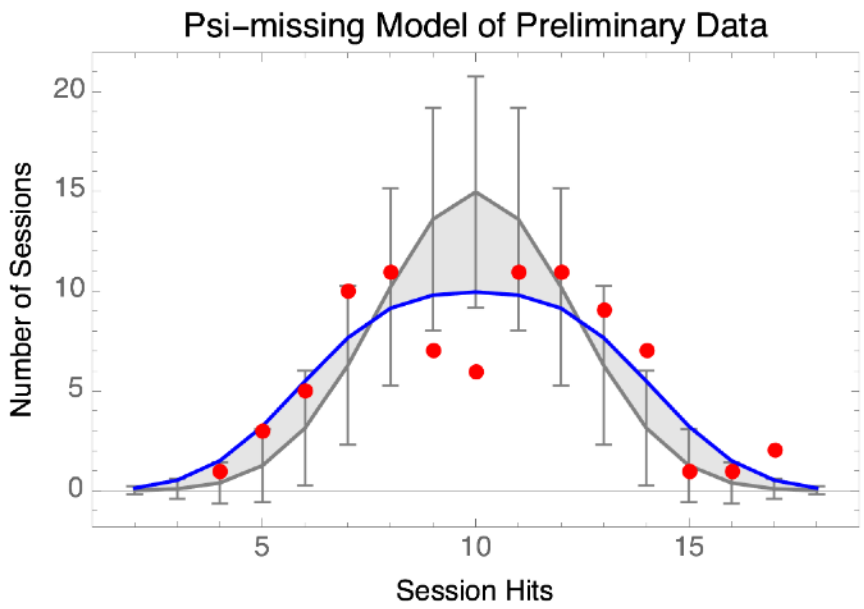
The model assumes a uniform effect size and adjusts the relative hitting/missing frequency, F, to fit the mean hit rate. The model's variance is controlled by an effect size parameter, D, which is defined as the absolute difference of the hit rate from 50%. Full details are given in the Appendix. The model (denoted as model M) gives a good description of the data when the frequency ratio is 1:1 and the hit rate is 60(40)% for hitting(missing), that is, F=1/2 and D=0.1. Figure 1 compares the model and the Null model of no psi effect. A Pearson chi-square distribution test (using Mathematica's PearsonChiSquareTest[] function), which is a common tool for comparing data to a theoretical distribution, finds that the preliminary data agrees with model M (p=.80), and rejects the Null model (p=.0014).

The good fit to model M indicates that the data are also consistent with a psi effect that is broadly distributed among sessions. A stronger statement in favor of model M is possible. Unlike the outlier model, model M pulls weight out of the distribution middle and reapportions it to the distribution's shoulder and wings (Figure 1). We specify an outlier model, O, which attributes 8 of 90 sessions to have extreme values in order to account for the high variance, and draws the remaining sessions from a Null model. A test of models M and O then compares the number of sessions in the distributions' centers – sessions with hits of 9, 10 or 11. The preliminary data has 26 sessions in this range. The expectations for model M, O and the full Null are 31.4, 41.0 and 44.7 sessions. The p-values for a count of 26 or fewer sessions for models M, O and N are, respectively (p=.139; p=.00064; p=.000045) so that both the outlier and Null models are strongly

rejected. Comparing models M and O directly, we find that the likelihood ratio of exactly 26 sessions with hits in the range 9-11 favors model M by about 115:1.

Figure 1.

The Distribution of Session Hits for 90 Preliminary Data Sessions.



Note: The blue trace is the expectation of the frequency of session hits for a model with equal parts of hitting and missing in alternate sessions with hit rates of 60% or 40%. The gray trace is the Null distribution curve (no psi effect). The red dots show the distribution of session hits in the preliminary data. Horizontal axis: number of hits per session; vertical axis: number of sessions with a given hit value. The error bars are 90% confidence intervals on the Null distribution. For clarity, error bars of model M are not shown, but they have approximately the same extent, and would be centered on the model's curve.

The psi-missing model M therefore provides a plausible description of the preliminary data. An alternate model where a few outlier sessions give the high variance is less plausible because it would not suppress the center of the distribution as seen in the data and model M.

Whatever the interpretation, we find that the effect appears for both cohorts. Of the 90 preliminary sessions, there were 47 Meditator sessions and 43 Open sessions. The variance tests are significant for both groups (Meditator: $\boldsymbol{X}2(47) = 77.6$; p = .0025; 400k MC iterations; Open: $\boldsymbol{X}2(43) = 73.2$; p = .0021; 400k MC iterations). Lastly, model M has an effect size (hit/miss rates of 60/40 percent) that is comparable to other reported psi effects. A 60% binary hit rate is roughly equivalent to a ganzfeld 4-choice hit rate of 33%. Meta-analyses of the ganzfeld give mean effects of around 32%, and subgroups of selected participants have hit rates as high as 40% (Baptista et al., 2016).

Assuming for the moment that our interpretation holds, it remains to explain why such a strong effect would not reproduce in the registered study. One possibility mentioned in the Introduction is that psychological factors changed during the two periods of data collection, and that these are necessary for an effect to appear. An alternate view is that psi declines mysteriously and its elusivity is beyond our control. The clearest response is that further studies are required to adjudicate the question, and it is precisely the purpose of Psi@Home to provide the needed data. In the meantime, our inclination is to favor a psychological explanation because we find it more parsimonious, at least as far as theoretical commitments are concerned. In fact, it is quite possible that participants were more motivated during the try-out sessions. The individual video calls with team members were meant to generate enthusiasm for the project, and care was taken to listen to the candidates' personal interests and emphasize the value of their participation in the research. The two try-out sessions were completed within days of the 45 minute call, when impressions from the video call were still fresh. Participants' positive attitudes and motivation for the registered study may have diminished because the study invitation arrived after a delay of 1-2 months and participants had no personal contact with team members before the study launch (a

brief group video call at the study's launch had a low attendance of about 20%). The requested task of 4-6 sessions was considerably more than the try-out and participants were under a deadline to finish. These factors contrast with the try-out period and may have been de-motivating. One can hypothesize that psi performance was weakened by a combination of stress and a lack of motivated engagement.

Another possibility is that mixtures of psi hitting and missing occurred within sessions for the registered data. The zero-mean variance test's power weakens if psi hitting or missing is not stable throughout a given session. In that case, even if the absolute strength of the psi effect is maintained, the test can fail to detect an effect. Tests sensitive to this eventuality which are based on autocorrelations are currently under study.

Conclusions

We have reported on a new platform for collaborative psi research with selected cohorts. The Psi@Home platform uses an downloadable application that allows people from the across the world to participate in experiments by doing sessions at home. The at-home design permits the establishment of cohorts with substantial numbers of participants and highly restrictive selection criteria. The platform employs a modular approach to experimental design that treats research teams as a fundamental element of experimental studies. It is envisioned as a "user facility" that external researchers can use to undertake psi studies with lower costs and faster execution. We hope this will make psi research more accessible to the scientific community. A study to compare two cohorts was easily completed within a month's time, validating our design goal of high data rates and reduced overhead for studies.

The study, which compared a cohort of experienced meditators with a general public cohort, yielded a non-significant result on 4 of 5 pre-registered statistical tests for a psi effect. A fifth test

was marginally significant, but confidence in its evidential value is diminished given the other negative results. On the other hand, a preliminary data set, whose collection was pre-registered as part of the process of recruiting cohort members, yielded a highly significant result for one of the two measures of an effect that we undertook for the formal study. The variance of session hit rates was nearly 4 standard deviations above null expectation, and in the direction expected for a model of psi-missing. We interpret this as evidence of a psi effect and show that a simple model of psi-missing accords well with the data's distribution of session hits. The absence of an similar effect in the formal study may be due to differences in participants' experiences between the two periods of data collection. While we feel that this interpretation is a plausible one, it remains a speculative proposal that is limited by the use of post-hoc analysis and modeling. A confirmation will need input from further study and data.

Authors' contributions

PAB designed and created the Psi@Home platform. All authors designed the study together. Cohort recruitment and data collection was done by JB and AB for the Open cohort; and by PAB for the Meditator cohort. PAB carried out the statistical analyses. PAB wrote the manuscript and MV helped with revisions.

Declaration of interests

The Authors declare that there is no conflict of interest. The study protocol was reviewed and accepted by the host institution's ethics committee. This work was generously supported by Grant 356/2020 from the Bial Foundation.

References

Bancel, P. A., (2018). Simulating Questionable Research Practices. In, *Proceedings of Presented Papers: The Parapsychology Association 61st Annual Convention*. https://doi.org/10.13140/RG.2.2.12941.64487

Bem, D. J., Tressoldi, P., Rabeyron, T., & Duggan, M. (2015). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. *F1000Research*, 4, 1188. https://doi.org/10.12688/f1000research.7177.2

Bierman, D. J., Spottiswoode, J. P., & Bijl, A. (2016). Testing for Questionable Research Practices in a Meta-Analysis: An Example from Experimental Parapsychology. *PLOS ONE*, 11(5), e0153049. https://doi.org/10.1371/journal.pone.0153049

Bösch, H., Steinkamp, F., & Boller, E. (2006). Examining psychokinesis: the interaction of human intention with random number generators--a meta-analysis. *Psychological Bulletin*, 132(4), 497.

Carpenter, J. C. (2004). First Sight: Part One, A Model of Psi and the Mind. The Journal of Parapsychology, 68(2), 217–254.

Jahn, R. G., Dunne, B. J., Nelson, R., Dobyns, Y. H., & Bradish, G. J. (2007). Correlations of random binary sequences with pre-stated operator intention: A review of a 12-year program. *Explore*, 3(3), 244–253.

Kekecs, Z., Palfi, B., Szaszi, B., Szecsi, P., Zrubka, M., Kovacs, M., … Aczel, B. (2023). Raising the value of research studies in psychological science by increasing the credibility of research reports: The transparent Psi project. *Royal Society Open Science*, 10(2), 191375. https://doi.org/10.1098/rsos.191375

Lawrence, T. (1993). Bringing home the sheep: A meta-analysis of sheep/goat experiments. In, *Proceedings of Presented Papers: The Parapsychology Association 36th Annual Convention*.

May, E. C., & Marwaha, S. B. (2018). *The Star Gate Archives: Reports of the United States Government Sponsored Psi Program, 1972-1995. Volume 2: Remote Viewing, 1985-1995* (Vol. 2). McFarland.

Palmer, J. (2015). Experimental methods in anomalous cognition and anomalous perturbation research. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (p. 49-62). McFarland.

Radin, D., & Pierce, A. (2015). Psi and Psychophysiology. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (p. 230-243) McFarland. https://psycnet.apa.org/record/2015-48721-017

Rhine, J. B. (1969). Psi-missing re-examined. The Journal of Parapsychology, 33(1), 1.

Roney-Dougal, S. M. (2015). Ariadne's Thread: Meditation and Psi. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (p. 125-138). McFarland & Company, Inc., Publishers.

Rouder, J. N., & Morey, R. D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin & Review*, 18(4), 682–689. https://doi.org/10.3758/s13423-011-0088-7

Schlitz, M., Bem, D., Marcusson-Clavertz, D., Cardena, E., Lyke, J., Grover, R., … Delorme, A. (2021). Two Replication Studies of a Time-Reversed (psi) Priming Task and the Role of Expectancy in Reaction Times. *Journal of Scientific Exploration*, 35(1), Article 1. https://doi.org/10.31275/20211903

Schlitz, M., Wiseman, R., Watt, C., & Radin, D. (2006). Of two minds: Sceptic-proponent collaboration within parapsychology. *British Journal of Psychology* (London, England: 1953), 97(Pt 3), 313–322. https://doi.org/10.1348/000712605X80704

Storm, L., & Ertel, S. (2001). Does psi exist? Comments on Milton and Wiseman's (1999) meta-analysis of Ganzfeld research. *Psychological Bulletin*, 127(3), 424–433. https://doi.org/10.1037/0033-2909.127.3.424

Storm, L., & Tressoldi, P. (2017). Gathering in more sheep and goats: A meta-analysis of forced-choice sheep-goat ESP studies, 1994-2015. *The Journal of the Society for Psychical Research*, 81(2), 79.

Storm, L., & Tressoldi, P. (2020). Meta-analysis of free-response studies 2009-2018: Assessing the noise-reduction model ten years on. *Journal of the Society for Psychical Research*, 84, 193–219.

Storm, L., & Tressoldi, P. (2023). Assessing 36 Years of the Forced Choice Design in Extra Sensory Perception Research: A Meta-Analysis, 1987 to 2022. *Journal of Scientific Exploration*, 37(3). https://doi.org/10.31275/20232967

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.

Timm, U. (1983). Statistische Selektionsfehler in der Parapsychologie und in anderen empirischen Wissenschaften. [Statistical selection errors in parapsychology and other empirical sciences.]. *Zeitschrift Für Parapsychologie Und Grenzgebiete Der Psychologie*, 25(3–4), 195–229.

Tressoldi, P. E., & Katz, D. (2023). Remote viewing: A 1974-2022 systematic review and meta-analysis. *Journal of Scientific Exploration*, 37(3), 467–489. https://doi.org/10.31275/20232931

Tressoldi, P. E., & Storm, L. (2024). Stage 2 Registered Report: Anomalous perception in a Ganzfeld condition - A meta-analysis of more than 40 years investigation. *F1000Research*, 10(234), 234. https://doi.org/10.12688/f1000research.51746.4

Varvoglis, M., & Bancel, P. A. (2015). Micro-psychokinesis. In E. Cardeña, J. Palmer, & D. Marcus-son-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 266–281). McFarland.

Varvoglis, M., Bancel, P. A., Bailly, J.-P., Boban, J., & Ahmed, D. S. (2019). The Selfield: Optimizing precognition research. *Journal of Parapsychology*, 83(1), 13–24. https://doi.org/10.30891/jopar.2019.01.02

Varvoglis, M., Bancel, P. A., Si Ahmed, D., Bailly, J. P., & Béguian, C. (2013). The sharefield: A novel approach for forced-choice GESP research. *Journal of Parapsychology*, 77, 181-183.

von Lucadou, W., Romer, H., & Walach, H. (2007). Synchronistic phenomena as entanglement correlations in generalized quantum theory. *Journal of Consciousness Studies*, 14(4), 50.

Walach, H., Kirmse, K. A., Sedlmeier, P., Vogt, H., Hinterberger, T., & Lucadou, W. von. (2022). Nailing Jelly: The Replication Problem Seems to Be Unsurmountable. Two Failed Replications of the Matrix Experiment. *Journal of Scientific Exploration*, 35(4). https://doi.org/10.31275/20212031

Watt, C. (2024). *Ganzfeld ESP research: Building on lessons learned*. Bial Foundation. https://fbial.yggycloud.com/getmedia.aspx?guid=b6223f751263faef0f2245bf00cc7ff8

Woods, J. H. (1927). The yoga system of Patanjali. Harvard University Press.

Appendix

Definitions of terms and the zero-mean variance test

The Null distribution of session hits is a binomial B[N, p], with N the number of trials and p=1/2 the binomial probability. Discrete binomial statistics approximate continuous gaussian statistics as N grows large and the correspondence is quite good when N=20 and above. In the text, we borrow a few terms from gaussian statistics and employ "z-score" for the number of excess hits divided by the theoretical binomial standard deviation:

$$z = 2\frac{(hits - \frac{n}{2})}{\sqrt{n}}.$$

Our usage should not be confused with a standard normal variable, which is the common meaning of z-score. We also refer to the sum of squares of N "z-scores" as the "zero-mean variance of N session z-scores":

$$\chi^2_N = \sum_i^N z_i^2 .$$

This is the quantity ***X***2(N) in the text. It is a discrete random variable that closely follows the (continuous) chi-square distribution with N degrees of freedom. A normalized value can be had by dividing ***X***2 by N, as is typically done for (theoretical) variances, but we prefer to cite the raw ***X***2 in this paper. "Zero-mean" signifies that ***X***2 is calculated about the theoretical mean z=0, instead of the sample mean (as is done for sample variances). This is preferred for our ***X***2 statistic since it allows a sensitivity to net psi-hitting (or missing) that would be lost if the variance were calculated relative to the sample mean.

Approximate p-values for ***X***2 can be estimated from the corresponding chi-square distribution.

We use more accurate Monte Carlo estimates of p-values when stating results. The MC procedure also allows p-values estimates when the number of trials per session varies.

Monte Carlo calculations

P-value estimates for the zero-mean variance are done by MC calculations on the *Mathematica* platform. A vector of z-scores for a study of K sessions is simulated using the RandomInteger[*dist, K*] function, where *dist* is the binomial distribution with N trials and probability p (p=1/2 for the Null distribution, but see below for other models). A value of $\boldsymbol{X}2$ is calculated from the z-score vector, and the process is iterated to give a simulated distribution for $\boldsymbol{X}2$. Empirical values from the experiment are then compared to the $\boldsymbol{X}2$ distribution to yield p-values. Note that the number of iterations can be increased to give a desired accuracy; we typically estimate p-values to an accuracy of better than 10%.

Difference test of cohort hit rates

The registered experiment gave Meditator and Open cohort hit rates of 48.8% and 51.8%, respectively (hits of 780 and 825; trial N's of 1597 and 1593). An effective z-score for the difference of cohort hit rates is given by :

$$z = \frac{HR_{Open} - HR_{Med}}{\sqrt{HR_{av}(1 - HR_{av})(\frac{1}{N_{Open}} + \frac{1}{N_{Med}})}}$$

where HR_{av} =.5031 is the weighted average hit rate, so that z=1.665 (p=.096, two-sided).

The psi-missing model M

Under the assumption of a psi effect, we ask if contributions to the high variance in the preliminary data come from the participant population as a whole, or only a few high performers. The question is important because the Psi@Home platform aims to elicit effects broadly.

Model M mixes hitting and missing sessions of uniform psi strength. The strength parameter, D, is the offset from a 50% hit rate and is defined on [0,1/2]. A parameter F, defined on [0,1], sets the proportion of hitting or missing sessions. The session hit rates are then 1/2±D, and the fractions of sessions with psi-hitting/missing are F and (1-F). Note that hits for the model

sessions are binomial variables, so that "hitting" sessions may produce hit rates less than 50%, and vice versa for "missing" sessions.

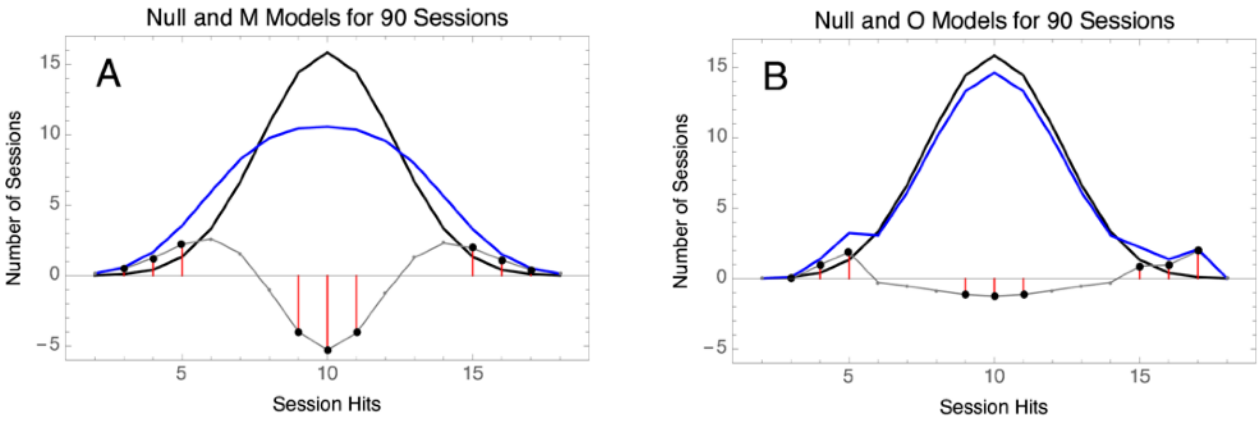
We adjust parameters D and F to give agreement with the mean and $\boldsymbol{X}2$ of the 90 preliminary sessions. It is fine to do this by inspection since we use the model M to draw comparative inferences, rather than determining precise parameter values. The preliminary data has a mean hit rate of HR=50.6% and $\boldsymbol{X}2$=150.7. Setting D=0.10 and F=1/2 gives MCE of (HR=50%; CI90(48.1, 51.9); and $\boldsymbol{X}2$=158.9; CI90(126.8, 193)), where the 90% confidence intervals are determined from MC. The parameter settings give a good fit to the data's mean and $\boldsymbol{X}2$.

We contrast model M and a model O, which assumes no psi effect except for 8 sessions with extreme hit rates. The outlier sessions are set to the empirical values of the 8 most extreme sessions in the preliminary data (4 sessions with 5 excess hits; and 2 sessions each with 6 and 7 excess hits). Model O yields (HR=50.0%; CI90(47.6, 52.4); and $\boldsymbol{X}2$=131.5; CI90(118.6, 159.7)), which is also consistent with the preliminary data.

Comparing M and O with the Null finds that the distribution tails are quite similar, and show a marked difference for the distribution centers (Figure A.1). Model M moves weight out of the center, which decreases the frequency of sessions with hits in the range from 9 to 11. A test of the session counts in this range therefore can distinguish between the broadly distributed effect of model M, and model O, which restricts an effect to a small number of sessions. P-values for obtaining 26 sessions in the center range (as found for the preliminary data) is calculated by MC for each model. The likelihood ratio for models M and O is had by estimating the probability of exactly 26 sessions occurring in the center range for each model by 400k MC iterations, and taking the ratio of frequencies.

Figure A.1

Comparison of Models M, O and Null



Note: The plots compare models M and O with the Null model. The horizontal axis is the number of hits in a session and the vertical axis is the number of sessions at the respective hit value, given a total of 90 sessions. Black traces: Null model; gray traces: differences of models M and O from the Null. Red bars highlight the differences at the distribution tails and centers. The distribution center is strongly suppressed relative to the Null for model M, and only slightly so for model O. Models M and O are distinguished by comparing the session counts in this range. The tails dominate the zero-mean variance, and do not allow a statistical discrimination of the models.

Effect size for model M

The strength parameter for the hit rate deviation is D≈.10 for model M. The effect size for a trial is

$$ES = \frac{D}{\sigma} = \frac{D}{\sqrt{p(1-p)}}.$$

For D=.1 and p=1/2 we have ES=.2 which is comparable to recent meta-analytic estimates for ganzfeld and remote viewing databases (Tressoldi & Katz, 2023). The D parameter can be converted to a Null offset for 4-choice protocols as the standard ganzfeld (G) if p=1/4. In this case G = .866*D = .087, which corresponds to a ganzfeld hit rate of 33.7%.