

# Blackmore-Sargent Controversy – A Reconsideration

Suspensions of cheating by Carl Sargent, a British experimental psi researcher in the 1970s, voiced by Susan Blackmore, a leading critic of parapsychology, continue to be presented by some as evidence against the existence of ESP. But close scrutiny fails to support Blackmore's claims, as revealed here by Professor Chris Roe, a psychologist and ESP experimenter at the University of Northampton and former president of the [Society for Psychical Research](#).

A version of this article is published in the Magazine of the Society for Psychical Research, issue 2, 2021.

## Background

Research using the [ganzfeld technique](#) remains one of the strongest lines of evidence for psi effects in the laboratory.<sup>1</sup> An extensive database of 113 studies comprising nearly 5,000 trials that have been carried out since 1974 demonstrates a small but robust and replicable effect.<sup>2</sup> Combined outcomes from experiments are reported in meta-analyses, but can be vulnerable to distortion if they include experiments that are regarded as suspect in some way.

In the case of ganzfeld research, some commentators have raised concerns about the validity of data reported by [Carl Sargent](#). For example, David Marks sets aside the evidence for an ESP effect in ganzfeld data on the grounds that 'the waters have been muddied, or poisoned even, by accusations of data manipulation and fraud. These same four meta-analyses yielding positive psi hit rates included the highly contentious studies conducted by Carl Sargent'.<sup>3</sup> The 'serious doubt' that Marks refers to derives from a [report by Susan Blackmore](#) following a visit to Sargent's Cambridge laboratory in November 1979 that alleges deviations from protocol and possible malpractice. This essay provides a summary and evaluation of these allegations.

After failing to produce evidence for psi in a ganzfeld experiment that formed part of her PhD,<sup>4</sup> Blackmore hoped to observe at close quarters a colleague who was using the same methods but with much greater success. There is an art to conducting research in the social sciences, and the kind of tacit knowledge that can facilitate success is, by definition, difficult to articulate and rarely features in published reports, so visits of this type can be extremely informative.

Blackmore stayed for eight days and was able to witness thirteen sessions, including six hits, well above the 25% expected by chance. She submitted a report to the Society for Psychical Research (SPR) concerning the visit (a condition of a grant she received for the purpose), in which she described Sargent's experimental approach but also raised concerns about discrepancies she had observed that she interpreted as evidence of cheating by him. She shared her report with Sargent and encouraged him to produce a response for publication. When it became clear that

he would not respond, she published a slightly emended version in the July 1987 issue of the *Journal for Psychical Research*.

## Sargent's Methodology

In order to understand the methods used by Sargent, we need to be mindful that in the 1980s many of the features of the ESP ganzfeld experiment (such as using a random method to choose the target image, providing the participant with a set of images for judging, and recording the outcome) was not automated but was managed by the researchers themselves. Such studies are now referred to as 'manual ganzfeld', to distinguish them from later 'autoganzfeld' designs in which key features are carried out by the computer. Even with the best of intentions, there is much greater scope for human error with a manual ganzfeld design.

Bem and Honorton's 1994 report on experiments conducted at Honorton's PRL laboratory signalled a shift in approach, and since then virtually all ganzfeld studies have been of the 'autoganzfeld' type. Experiments typically involved a 'sender' or 'agent', who was provided with a randomly selected target image to concentrate on, and a 'receiver' or 'subject' who reported their impressions during the ganzfeld period (termed a 'mentation') with the intention that they would relate to the target image and enable them to identify it. The subject would subsequently be presented with the target image and three decoys by their experimenter (who also did not know which was the target), and these would be rated or rank ordered based on similarity to their mentation.

Blackmore's report is balanced and transparent. She acknowledges that the laboratory environment was qualitatively different from hers in ways that could have been psi conducive, particularly in creating a relaxed and supportive atmosphere and cultivating confidence in success. These qualities have been commonly associated with experimenter effects in parapsychology.<sup>5</sup> She also concedes that the design 'seemed to exclude very efficiently the possibility of sensory leakage'. She continues, 'Duplicate target sets were used so that no handling cues were available. The subject and subject's experimenter were entirely isolated from the agent ... In the sessions I observed I could see no means of sensory leakage unless protocol were violated. I observed no such violations of protocol at this stage'.<sup>6</sup> However, she did have concerns about the randomization method and its potential to allow for cheating.

Sargent's team used a database of 108 pictures arranged in 27 sets of four, such that the images in any particular set were as distinct from one another as possible (so that they could be easily discriminated between during judging). There were two copies of each picture: one copy was placed in an individual sealed envelope to be provided to the agent as the target; the other copy was included in an envelope with the other pictures of its set to be used by the subject during judging. This is an important feature, since it ensures that there will be no handling marks on the picture the agent has been concentrating on that might indicate to the subject that it was the target (Blackmore's own ganzfeld study did not have duplicate images for judging).

Selecting the target needs to involve a method that ensures the subject (and indeed their experimenter) can't second-guess what it might be; for example, if we avoided pictures with children because yesterday's target was a children's party. At Cambridge, a researcher who does not interact with the subject, called the Randomizer, would consult published tables of random numbers and the first digit that fell in the range 1-28 (13 wasn't used) determined which picture set to use for that trial. Which image to use as the target within that set was decided by taking a pile of 20 small envelopes (5 containing As, 5 with Bs, etc.), cutting the deck and then counting down to the nth envelope, where n was the first number between 1 and 20 in the table of random numbers.<sup>7</sup> The agent would take the four large envelopes that belonged to the set selected for that trial, along with the small envelope that told them which of the large envelopes, A-D, to open. They only opened the small envelope once they were secure in the agent's room and the trial had begun, and would then open the large envelope indicated to reveal just that picture.

Meanwhile, the participant would experience ganzfeld stimulation and give an ongoing commentary while being monitored by the session experimenter. Once the ganzfeld period was over, the experimenter would go to the office to retrieve the judging envelope that had been left out for them and they would present the participant with all four images that it contained for judging against their impressions. After the trial, the small envelope wasn't returned to the deck; instead, a replacement envelope was added. These envelopes were already sealed and unmarked, so were kept in separate drawers so that the correct letter could be added after the trial to ensure there were equal numbers of each in the deck for the next trial (that is, if the target for that trial had been a 'C' then an envelope from the 'C' drawer would be added to the deck).

## **Blackmore's Observations and Speculations**

Blackmore speculated about possible mechanisms for manipulating randomization, focusing on the possibility that the deck of small envelopes might be biased rather than having equal numbers of the four options, A-D, so that some targets would be more likely than others. Alternatively, the target letter for a particular session might be taken from one of the drawers rather than from the shuffled deck, so that its identity would be known with certainty.

She focused on trial 9 in the sequence that she observed, after discovering that one of the envelopes in the B drawer was missing; the target for that trial turned out to be picture B. Although Sargent was not formally involved in the trial, she noted that he had stepped in to act as randomizer and was then present during the judging phase and 'seemed to push the subject towards picture B'. That would be completely inappropriate if he had any way of knowing which letter had been chosen as the target image in this set. Of course, the small envelopes are sealed when the randomizer selects one. But if the randomizer knew the deck was biased or had in fact selected the envelope from a drawer, then this breach of security would be enough to void the session. The subject ranked picture 'B' as the number 1 selection, and this turned out to be the target image, so the trial was recorded as a 'hit'.

Blackmore disclosed her concerns to Trevor Harley (Sargent's co-experimenter). They checked all the envelopes in the deck to see if each letter was equally represented, as they should be. However, they discovered an extra 'A' and 'B' and one fewer 'C' and 'D' than expected. While clearly this deviation from expectation should not have happened at all, it is much too small a shift to account for the study outcome – Sargent estimated that it might give a 3% advantage, when the overall hit rate was 42%) – and so provides only meagre evidence in support of Blackmore's first mechanism. If the method for cheating was to take a small envelope directly from one of the drawers (or from elsewhere in the room) to give to the agent so that the target image was known for certain, then it's not clear why there would be any need for the distribution of letters in the deck to be biased.

Harley and Blackmore also conducted a search of the office, looking for a lab book that was used to record details of each session. The book was not found, but they did discover some additional small envelopes in various places (under papers, in another drawer). These were sealed, and when opened revealed a 'C', a 'D' and a batch of 3 'A's. Again, this is unexpected, but their haphazard placement is very difficult to fathom, or to align with the other observations – if known letters were secreted about the office then why should drawer numbers fall and why should the shuffled deck have uneven numbers? It would have been much more straightforward to simply have spare cards and open envelopes, and write whichever letter was needed and seal it in the envelope; this would have taken no more time than hunting around for whichever secret location contains the 'correct' envelopes. For a supposedly astute investigator such as Sargent, the supposed method seems very naïve indeed.

## **Blackmore's Predictions of Cheating Questioned**

Blackmore went on to make some predictions based on her suspicions, most of which sadly remain untested. However, she did reasonably predict 'if one person were cheating, the most significant results should occur when they were acting as agent or experimenter', and she claimed that indeed scores were higher in sessions that involved Sargent in one of these roles, both among the trials she observed – most of which formed part of experiments later reported by Sargent, Harley, Lane and Radcliffe (1981)[8](#) and Sargent and Matthews (1982),[9](#) and also in data reported by Ashton, Dear, Harley and Sargent (1981).[10](#)

The 1982 study involved only two researchers, so Sargent would always have had to take the role of either agent or experimenter, making it impossible to test Blackmore's prediction. That experiment gave a significant 46% hit rate across 26 trials, but none of the internal effects (differences between conditions or between different participant groups) predicted by Sargent and Matthews were supported, which would be a missed opportunity for a determined cheat whose aim would surely be to confirm those predictions.

The purpose of the 1981 study was also to test for internal effects, specifically whether experienced subjects would perform better than naïve ones, and whether a 30-minute period of ganzfeld stimulation would be more effective than a 15-minute period. Overall scoring in the experiment was at chance level, which again seems

surprising for a study that is alleged to have involved cheating. There was no advantage for the longer ganzfeld period over the shorter one, failing to confirm the authors' hypothesis, but there was suggestive evidence for a within-session incline effect (impressions reported later in the session were more accurate than those described earlier).

However, sessions in which Sargent was experimenter (a role in which he might be able to influence the subject's selection during judging) ran opposite to trend, so much so that the analysis would have been significant without his contribution. If Sargent were cheating, it is perplexing that he would use this to diminish the very effects that the experiment was designed to detect.

For the Ashton et al. study, Blackmore's claim is demonstrably false, and suggests she misunderstood data presented in that paper. Here, the four co-authors took turns to play the respective roles of experimenter, subject and agent. They report that the subject who achieved the greatest number of hits was Sargent, occupying the role in which he had least opportunity to affect the target selection or to bias judging. For the small number of sessions that Blackmore observed, Sargent served as randomizer on five trials, of which four resulted in hits, and as experimenter for four sessions, of which two were hits (Matthews was randomizer for just two trials but both were hits).

## **Sargent's Response to Blackmore**

Sargent's co-researchers Trevor Harley and Gerald Matthews published a joint response to Blackmore's allegations,[11](#) taking issue with the 'cheating hypothesis' on the grounds that it was so vague that any deviation from the standard procedure would be deemed a confirmation of it rather than attributable to honest human error. A more focused hypothesis regarding cheating would lead to predictions of a particular pattern of behaviours rather than the random set of anomalies that were observed.

They also argue that a reasonable explanation for some of the errors had been offered before the fact by Sargent; for example, that the missing envelope in the 'B' drawer was removed because it had a bent corner that might identify it. Blackmore recently questioned whether this represented sufficient grounds for removing it, 'because a bent corner could not affect the choice of target', but researchers knew this envelope came from the 'B' drawer, so that if it were ever to find its way into the deck then the experimenter performing the randomization would immediately know what letter it contained. I have no doubt that if such an envelope had been retained then it would have been cited as a design flaw that allowed for some forms of cheating.

Harley and Matthews also note that since experiments were concerned with exploring the effects of personality or situational variables upon the outcome, rather than demonstrating (yet again) that above-chance scoring could occur, experimenters would not know at the time which trials were 'supposed' to be successful. From this perspective, the suspect trial 9 was not expected to produce a hit since it involved an introverted subject in a short-duration session. Cheating in order to produce a spurious hit would amount to self-sabotage. With regard to

Sargent 'pushing' the subject toward picture 'B', Harley and Matthews add that when the subject's impressions were given to an independent judge, they gave picture 'B' an even higher rating, even without the pushing from Sargent.

Finally, they consider the claim that because the deck had an uneven number of As, Bs, etc., this could lead to dubious hits if the experimenter were able to push the subject to select in the direction of the bias. This hinges on any bias in the deck being realized in terms of the actual targets selected. Harley and Matthews present all the target selections for the Sargent and Matthews series, which comprised 7 As, 7 Bs, 5 Cs and 7 Ds; that is, almost exactly what one would expect from a chance distribution. They conclude, 'Blackmore's report is loaded with prejudicial reporting and inconsistencies. If analysed correctly, the data clearly show that her observations are best explained by a "random errors" hypothesis'.<sup>12</sup>

Sargent also published a response,<sup>13</sup> claiming that Blackmore had herself given the material from the suspect trial 9 to an independent judge but that the target picture was still correctly identified; he also stated that an account of this given by her in her original report had been omitted from the published version. Blackmore confirms this in a reply to Sargent,<sup>14</sup> in which she includes the following direct quote from that original report, 'I asked another person in the lab (DG) to judge the pictures against S's protocol. He, narrowly, gave B rank 1 confirming the original choice. However, he was not sure that he had not previously heard talk about this particular session'. It seems odd that Blackmore did not confirm that DG was entirely ignorant with regard to trial 9 before conducting the time-consuming re-judging. That leaves her open to speculation that, if the independent judging *had* failed to identify picture B as the target, then she might not have gone on to explore whether DG had possibly been exposed to that previous talk.

Sargent also describes another session (7) in which he, as experimenter was concerned that the participant was neglecting one of the pictures in the judging set, and repeatedly encouraged the participant to consider correspondences between it and their mentation – so much so that Blackmore had asked outright whether Sargent thought the neglected picture was the target. Ultimately, it turned out that it was not the target, and in fact the subject correctly chose one of the others as the first choice. In other words, if Sargent's 'pushing' had been successful, it would have prevented a hit rather than created one.

It is disappointing that this case is not described by Blackmore, when, according to Sargent, it very clearly was interpreted by her at the time as an instance of pushing. Blackmore reports<sup>15</sup> that her own notes do not refer to any bias on Sargent's part in session 7, though they do confirm that he thought D was the target whereas the subject ranked A as first choice and got a direct hit.

Having myself been involved in over 300 ganzfeld trials, I think it is reasonable that the session experimenter's role might extend to encouraging the subject to consider all of their impressions and to ensuring they are not overly influenced by elements of their experience that might be an artefact of the ganzfeld procedure (such as sensations of floating or falling). To someone unfamiliar with the method, this could seem like 'pushing', even though the experimenter would of course have no idea which of the pictures was the designated target. In any case, the efficacy of



this kind of subtle social pressure has to be evaluated in the context of striking qualitative correspondences between ganzfeld imagery and target features (reported, for example, by Ashton et al.).

## Parapsychological Association Investigation

Blackmore's reported suspicions led to an investigation in 1984 by the [Parapsychological Association](#) (PA). A committee chaired by Dr Martin Johnson made a brief report concluding there was not sufficient evidence to support the charge that Sargent's experimental procedures were 'unethical'. However, it did reprimand him for improper behaviour in having failed to respond to the charges made against him despite repeated requests to do so.

In a letter to *Skeptical Inquirer*<sup>16</sup> Blackmore states that the PA committee 'deemed it indefensible of me to go into Sargent's office, open his randomization envelopes, and set traps – all of which I knew was necessary for me to try to get to the truth, and Sargent had explicitly given me the run of his office'. This is inconsistent with the PA's actual report, which states that they 'felt that Blackmore's use of covert maneuvers to obtain information about possible fraud may or may not be considered unethical, depending on the premises upon which one makes such a judgment'. Any 'reprimanding' of Blackmore was for 'making her report essentially confidential and in then apparently allowing its contents to be "leaked" ' (p. 3).

Following the controversy, Sargent allowed his professional membership of the PA to lapse. He stopped conducting original research soon after, because he felt that his work had been tarnished by the controversy and also because a change in leadership at Cambridge made his interest in parapsychology untenable.<sup>17</sup> His last contribution to the field was to co-author with Hans Eysenck the introduction to parapsychology, *Explaining the Unexplained: Mysteries of the Paranormal*.

## Conclusion and Implications

It is clear from this brief review that Blackmore's concerns about Sargent's research practices are technical and quite complex. Her proposed mechanisms for cheating rely on the convergence and interplay of rather subtle factors. One might be to purposely choose a picture that is expected to align with the subject's personal preferences; another might be to know the target choice and exert pressure during judging to override overt similarities between the subject's ganzfeld impressions and the imagery found in the four pictures.

In my judgement, the evidence that is presented in support of Blackmore's suspicions is circumstantial, requires interpretation, is based on a very small number of observations, and depends on maintaining a number of alternative (and potentially contradictory) cheating scenarios simultaneously. So, while it is indeed disappointing to find deviations from the high standards of practice we must maintain in parapsychological research, it is difficult to see how one could confidently conclude from this scattergun approach that cheating had been demonstrated.

To an extent, this complexity and subtlety of the case is appreciated in the exchange that was published in the SPR's Journal in 1987. Unfortunately, more recent accounts of the controversy tend to be much briefer and more polemical, as found for example in David Marks's comments referred to earlier. When Blackmore revisited the case in a 2017 *Skeptical Inquirer* article,<sup>18</sup> she wrote, 'It became clear that Sargent had deliberately violated his own protocols and in one trial had almost certainly cheated.' Such a definitive conclusion is clearly not justified by the material reviewed here. Blackmore goes on, 'It matters that Sargent's experiments were seriously flawed. It matters that Bem included these data in his meta-analysis without referencing the doubt cast on them. It matters because Bem's continued claims mislead a willing public into believing that there is reputable scientific evidence for ESP in the ganzfeld when there is not.'

Setting aside the fact that the Bem and Honorton meta-analysis does *not* include Sargent's experiments (it is concerned only in reporting Honorton's own autoganzfeld experiments), it is quite astonishing for Blackmore to extrapolate from her suspicions about practices she observed at one laboratory during an 8-day visit to justify to dismiss 35 years of research carried out by 46 different principal investigators. About 65% of those studies are of the 'autoganzfeld' type to which concerns about randomization and target selection cannot, by definition, apply. Recent experiments show no indication of a decline,<sup>19</sup> and are not dependent on the particular success of Sargent's (or Honorton's) laboratory. This means that a new meta-analysis that excluded their work would still be highly significant. However, on the basis of the material reviewed here, there are absolutely no grounds for creating this.

Chris Roe

## Literature

Ashton, H.T., Dear, P.R., Harley, T.A., & Sargent, C.L. (1981) A four-subject study of psi in the ganzfeld. *Journal of the Society for Psychical Research* 51, 12-21.

Blackmore, S. (1987a). A report of a visit to Carl Sargent's laboratory, *Journal of the Society for Psychical Research* 54, 186-98.

Blackmore, S. (1987b). (A response to Harley, Matthews and Sargent). *Journal of the Society for Psychical Research* 54, 275-76.

Blackmore, S. (2018). [Daryl Bem and psi in the ganzfeld](#). *Skeptical Inquirer* 42/1.

Blackmore, S. (2021). Letter to the editor. *Skeptical Inquirer* (March/April).

Harley, T., & Matthews, G. (1987) Cheating, psi, and the appliance of science: A reply to Blackmore. *Journal of the Society for Psychical Research* 54, 199-207.

Harley, T. (2019). Obituary: Carl Lyndwood Sargent. *Journal of the Society for Psychical Research* 83/2.

Marks, D.F. (2020). *Psychology and the Paranormal: Exploring Anomalous Experience*. London: Sage Publications.



Roe, C.A. (2016). Experimenter as subject: What can we learn from the experimenter effect? *Mindfield* 8/3, 89-97.

Sargent, C.L., Harley, T. A., Lane, J., & Radcliffe, K. (1981). Ganzfeld psi optimization in relation to session duration. *Research in Parapsychology* 1980, 82-84.

Sargent, C.L., & Matthews, G. (1982). Ganzfeld GESP performance in variable duration testing. *Journal of Parapsychology*, 159-160

Sargent, C. (1987). Sceptical fairytales from Bristol. *Journal of the Society for Psychical Research* 54, 208-18.

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.

Tressoldi, P., & Storm, L. (2021). [Stage 2 Registered Report: Anomalous perception in a Ganzfeld condition - A meta-analysis of more than 40 years investigation.](#)

## Endnotes

### Footnotes

- [1.](#) Storm & Tressoldi (2020).
- [2.](#) Tressoldi & Storm (2021).
- [3.](#) Marks (2020), 130-31.
- [4.](#) This study remains unpublished but consisted of 36 sessions. No direct hits data are reported and no raw data are provided, but the overall sum of ranks for the originally planned 20 sessions was 50, which is exactly what would be expected for a chance distribution (mean rank 2.500, CRz = 0.1). For an additional 16 trials, SoR is slightly better at 35 (mean rank 2.188, CRz = 1.0). Only one target set was used for both sending and judging, which leaves the study susceptible to leakage problems given that the experimenter had to collect the pictures for judging from the agent, and that the target picture may have displayed signs of handling during sending. The randomization was initially achieved by manual shuffling by the agent which would be classed as inadequate, but later involved random number tables. Participants were able to contribute more than one session which is problematic in that it introduces a self-selection bias and also compromises the assumption of data independence needed for statistical analysis.
- [5.](#) Roe (2016).
- [6.](#) Blackmore (1987a), 190.
- [7.](#) This description is from Ashton et al. (1981). Blackmore's description is more convoluted, but may be based on her personal observations.
- [8.](#) Sargent et al. (1981).
- [9.](#) Sargent & Matthews (1982).
- [10.](#) Ashton et al. (1981).

- [11.](#) Harley & Matthews (1987).
- [12.](#) Harley & Matthews (1987), 205).
- [13.](#) Sargent (1987).
- [14.](#) Blackmore (1987b).
- [15.](#) Blackmore (1987b), 276.
- [16.](#) Blackmore (2021), 63.
- [17.](#) Harley (2019).
- [18.](#) Blackmore (2018).
- [19.](#) Tressoldi & Storm (2021).

© Psi Encyclopedia