A REPORT OF A VISIT TO CARL SARGENT'S LABORATORY*

by Susan Blackmore

The report which we publish below is based on the 1979 report deposited in typescript in the Society's library where it was available to interested persons on request. The author has now decided that it would be fairer to all concerned if this report were now brought into the public domain. Accordingly we are now publishing it in a slightly emended form and have invited those who consider themselves to be explicitly criticized therein to write a rejoinder.—Editor.

ABSTRACT

In 1979 I visited the laboratory of Dr. Carl Sargent at the University of Cambridge, to observe highly successful ganzfeld psi experiments then in progress. I observed 13 sessions, of which six were direct hits. I considered whether the results might be accounted for by sensory leakage, experimental error, cheating or psi. I made observations of the sessions to test these hypotheses. The experimental design effectively ruled out sensory leakage. However, I observed several errors in the way the protocol was observed. Most of these occurred in the cumbersome randomisation procedure. It was not clear how these errors came about. Their origin might have been clarified by either (a) a statement from Sargent or his colleagues, or (b) by reanalyses of the raw data. However neither has been made available. Sargent's nine ganzfeld studies form a considerable proportion of the total ganzfeld database. In view of Sargent's unwillingness to explain the errors found, or to make his data available to other researchers, I suggest that these results should be viewed with caution.

Introduction

In November 1979 I went to visit Carl Sargent's laboratory at the University of Cambridge. He had carried out numerous ganzfeld experiments with highly successful results (Sargent 1980). Meanwhile I had been unsuccessful in superficially similar ganzfeld experiments at the University of Surrey (Blackmore 1980).

The objective of the visit was to observe the methods and conditions used at Cambridge and compare them with those used at Surrey, to see whether any reason for the discrepancy in the results could be determined. Because of the possibility of a psi-mediated experimenter effect, Sargent and I hoped to carry out experiments in which we would both act as experimenter while using the same subjects and procedure. Sargent kindly invited me to visit his laboratory for a month. The Society for Psychical Research (SPR) provided a grant to cover my expenses while there. In the event I was only able to stay eight days from November 22–30 1979.

During the visit I observed several errors in the way that the protocol was observed and the randomisation procedure carried out. The source of these errors was unknown. After the visit I wrote a report for the SPR (a condition of

* Acknowledgements

I wish to thank Dr. Carl Sargent for inviting me to visit his laboratory and the Society for Psychical Research for financial support.

the grant) which was placed in the Society's office and was available to any member who wished to see it.

The account which follows is based on four sources of information. 1. My original report for the SPR (which is still available from them). 2. My notes which I made during the visit to Cambridge. 3. My private diaries written each day. 4. Letters between myself, Sargent and other interested parties.

The Experimental Procedure

At the time of my visit three ganzfeld experiments were in progress. I observed a total of 13 sessions. I either watched the experimenter and subject, or the agent, or acted as one of these myself. The experimenters were Sargent, Trevor Harley and student experimenters (G. M., J. L. and K. R.). The subjects and agents were all of these, plus other students and friends of the experimenters.

All experiments used the same procedure, outlined below, with the following variations. In one experiment subjects could remain in ganzfeld for as long as they wished. In another, sessions lasted either 15 or 30 minutes, and the third involved a study of subject–agent pairs. There was also one session conducted at a private house for the benefit of the BBC and for this the procedure was, necessarily, slightly different. With these variations, the procedure was as follows.

The subject arrived at the experimental room where Sargent or the student experimenter gave them coffee, chatted with them and, if the subject was a novice, explained the purpose of the experiment and the procedure. There was often music playing and the atmosphere was very informal and relaxed. In some cases the subject brought a friend to be agent, but in most cases an experimenter acted as agent.

When the subject was ready the experimenter gave him or her a pre-session questionnaire to complete. The subject then lay on a comfortable mattress on the floor and was prepared for the ganzfeld. Half ping-pong balls were fixed over the eyes with sellotape and cotton wool and white noise was played through headphones, adjusted to be comfortably loud. A red light was shone on the ping-pong balls. The subject was then left alone and the door shut. The experimenter's and agent's watches were synchronised from the start of the ganzfeld session.

The experimenter then retired to the control room from which he could watch the subject throughout the session through a one-way mirror. A microphone near the subject's head picked up everything that was said. This was relayed to the control room and was both recorded on tape and written down by the experimenter.

Meanwhile the agent alone (if one of the experimenters was to be agent) or the agent with an agent's experimenter, went along the corridor into Sargent's office to select the target for that session. There were 27 sets of pictures, each containing four black and white or coloured pictures, chosen by Sargent and Harley to be as different as possible from each other. One of these was selected by using random number tables. There were two copies of each set. One contained the four pictures in individual sealed large envelopes, for the agent. The other, duplicate set, had all four pictures in one envelope and this was left in the office. In each set the pictures were lettered A to D.

Next a small sealed envelope containing a letter A–D was selected and used to determine which of the four pictures in that set was to be target. The randomisation procedure is outlined in more detail below.

The agent took the four large envelopes and the small envelope (all still sealed) to a different building and into a soundproof booth. At a pre-arranged time (depending on the experiment) the small envelope was opened. This contained a letter A–D. The corresponding large envelope was then opened and the agent took out the picture and looked at it for the prescribed length of time, making notes on a sheet provided. He or she retained the small envelope with its letter. The other three large envelopes remained sealed. Afterwards he or she waited near a telephone in another room on that floor of the building.

At the end of the ganzfeld session the experimenter went into the subject's room, turned off the white noise, removed the headphones and ping-pong balls and gave the subject a post-session questionnaire to complete. He then went into the office and collected the duplicate set of pictures, left there by the agent. He laid them out in order in front of the subject and then went through the transcript of everything the subject had said. Each picture was marked, by the experimenter and subject together, on a scale of 0 to 2 for correspondence with each item of the transcript.

The various experimenters differed somewhat in their approach to the judging and in the extent to which they encouraged or guided the subject, but in all cases the total score for each picture was added up and the subject then asked to rank and rate (on a scale of 1–100) all four pictures.

Once the ranks and ratings were recorded the experimenter telephoned the agent and asked him to come over. He always used the same words when ringing. When the agent (and agent's experimenter when applicable) arrived they disclosed which picture was target and showed this, together with the other unopened envelopes, and the letter A–D, to the experimenter and subject. The rank allocated to the target was then known and a z-score based on the ratings was calculated.

The Randomisation

The randomisation procedure is briefly described in Ashton, Dear, Harley & Sargent 1981 and Sargent 1980. It was rather complex and I shall therefore describe it in more detail.

There were 27 sets of four pictures, numbered 1–12 and 14–28. First the agent selected one of these by taking an arbitrary starting point into the RAND random number tables, and taking the first number between 01 and 28 (excluding 13). This determined which set was to be used.

The pictures in each set were lettered A–D. Which was to be target was determined as follows. There was a pile of 20 small brown sealed envelopes constantly on the desk in the office. Each contained two pieces of white card enclosing a slip of paper bearing one of the letters A, B, C or D. There were five of each letter in the pile, the envelopes being all of the same type and unmarked. The agent (or agent's experimenter) opened the book of random digits, arbitrarily selected an entry point and took the first number between 01 and 20. He counted this number of envelopes down the pile and cut it. He then took the next number in the list, and counted down the pile again, taking the envelope indicated. Once in the soundproof room he opened this envelope and used the letter it contained to determine which of the large envelopes would be opened.

Afterwards the pile obviously contained only 19 envelopes. To restore it to 20 the one used had to be replaced; and by one of the same letter. In four drawers adjacent to the desk, spare envelopes were kept: A's in the top drawer, down to D's in the bottom drawer. Like those in the original pile they were, of course, of the same type and unmarked. Their contents were only known by which drawer they came from. After each session the experimenter looked up which letter had been used for that session, took an envelope from the corresponding drawer and placed it in the main pile. In this way the main pile could retain its contents unchanged.

Experimental Results

During my visit I observed 13 sessions. Of the 12 conducted at the laboratory, six were direct hits. This is a hit rate of 50 per cent when 25 per cent is expected by chance. Obviously the number of sessions was small but the results seemed to confirm Sargent's previous high rate of scoring.

Table 1. Summary of Sessions Observed

No.	*Expt*	*Subject*	*Experi-menter*	*Agent*	*Agent's Expter*	*Random-iser*	*Rank*	*S.B. Observed*
1	1	friend of A	G.M.	T.H.	–	C.S.	1	A
2	BBC	BBC	C.S.	D.G.	T.H.	S.B.	3	A
3	1	M.	G.M.	C.S.	–	C.S.	1	E
4	2	G.	T.H.	J.L.	–	J.L.	4	E
5	1	J.B.	C.S.	G.M.	–	G.M.	1	E
6	2	–	C.S.	J.L.	–	J.L.	2	A
7	1	J.	C.S.	G.M.	–	G.M.	1	E
8	1	H.A.	G.M.	C.S.	–	C.S.	2	E
9	2	R.	K.R.	T.H.	–	C.S.	1	E
10	2	S.B.	H.A.	R.P.	C.S.	C.S.	1	S
11	2	J.A.	T.H.	T.B.	K.R.	K.R.	3	Judging only
12	3	A.J.	S.B.	N.C.	T.H.	T.H.	4	E
13	1	S.B.	G.M.	T.H.	–	T.H.	3	S

Experiments 1. Sargent and Matthews 1982.
2. Sargent, Harley, Lane and Radcliffe 1981.
3. A study of Subject-Agent pairs.

Overall Sum of Ranks 27 $p = 0{\cdot}22$ (2-tailed).
Excluding BBC session 24 $p = 0{\cdot}16$ (2-tailed).

Observations Phase 1. Days 1–3

During the first three days of my visit I observed five sessions (not counting the BBC session). During these sessions I did not take part but just watched either the experimenter or the agent. I took detailed notes, intending to compare the procedure with my own. Of the five sessions, three produced a rank 1 or direct hit, one a rank 2 and one a rank 4. That is a hit rate of 60 per cent when 25 per cent is expected by chance (p = ·12). These results seemed quite unlike my own chance results.

The whole purpose of the visit was to try to determine the reason for the difference in results between my experiments and those of Sargent. I considered the following five hypotheses and made observations accordingly.

1. Differences in 'atmosphere'

Sargent's setting and procedure appeared to be potentially far more psi-conducive than mine. The room was much larger and more pleasant. There was music and coffee and the whole environment was much less like that of a laboratory.

2. Differences in experimenter

It was quite clear that the main experimenter was extremely confident about the expected results and conveyed this confidence to the subjects. The experimenter's role during the judging was also much more active with the experimenters, especially Sargent, often encouraging the subject, making suggestions and pointing out correspondences. This would allow for more influence by the experimenter, which might be good or bad. With a skilled experimenter, it might maximise the use of the available information.

3. Sensory Leakage

The design seemed to exclude very efficiently the possibility of sensory leakage. Duplicate target sets were used so that no handling cues were available. The subject and subject's experimenter were entirely isolated from the agent, from the time the watches were set until the phone call was made. By this time the subject had made his choice. 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.

4. Errors

Three questions arise here. First does the procedure allow for errors to take place? Second are those errors likely to be important to the results, and third did any errors actually occur?

The sort of accidental errors which might occur include incorrect replacement of pictures in envelopes, errors in the timing, in giving the right questionnaires, in the addition of marks or the calculation of z-scores.

First, the complex randomization procedure seemed to allow for errors to take place reasonably easily. For example, if an envelope were incorrectly replaced in the pile this would lead to a bias in the pile which might never be detected. Second, however, such a bias would produce only a small effect on the overall scores.

Third, only one error was observed during this stage. On one occasion, when the duplicate set of pictures was brought in for judging, it was found to contain only three pictures instead of four. The problem was efficiently resolved. J. L. (a student) rang the agent (G. M.) (in my presence) and asked him to see whether he had an extra picture by mistake and if so, to place it on the ground floor of the other building and then return to his place by the phone. Sargent then went to fetch it and the judging proceeded as usual. This sort of error can easily arise in experiments of this complexity but, if handled correctly like this, could not produce spurious results. No other errors were observed at this stage.

5. Cheating

I had no reason to suppose that anyone might be cheating. However, parapsychology is still a controversial subject and it is conventional to consider whether a protocol is proof against obvious methods of cheating, even though completely cheat-proof designs are not to be expected and are probably unattainable. My intention was to look for any obvious methods and to ensure that they were not taking place. In this way I could be reasonably certain that the only remaining hypothesis was that of ESP.

Before the visit I had thought of several possible methods. These involved the experimenter finding out which picture was target, and pushing the subject towards it, or the agent opening a different picture from the one specified by the randomization. I could now see that the experimental design made any of these methods extremely difficult. However, the complex randomization procedure seemed to allow for several methods of cheating. The observations necessary to check up on these were simple and unobtrusive and I believed them to be necessary if I was to convince myself and others of the validity of the results. This led to the second phase of the observations during which I checked various new hypotheses.

Observations. Phase 2. Days 4–5

During these two days I observed a further five sessions. I was subject in one of them. There were three direct hits, one rank 2 and one rank 3. This is a hit rate of 60 per cent, where 25 per cent is expected by chance. For these few sessions alone, the results are almost significant ($p = \cdot 055$). For the ten sessions observed so far the sum of ranks was 17 ($p = \cdot 016$).

Clearly chance was very unlikely to account for these results. Sensory leakage or simple experimental error had been excluded and so the remaining possibilities seemed to be either ESP or cheating. I should point out that it was probably clear to everyone in the lab that I was sceptical about the possibility of ESP. I believe that having a sceptical observer there was not particularly pleasant, but on the other hand the sessions proceeded in a relaxed and pleasant atmosphere and the results were not adversely affected by my presence. I did not tell anyone about the specific hypotheses I had in mind. I hoped only to make some simple observations which would exclude them to my own satisfaction.

With this in mind I considered whether any simple methods of cheating were possible within this experimental design. I considered the following hypothetical methods and ways of detecting them.

1. The pile of small envelopes could be biased. The experimenter would then know which picture would be target and could 'push' the subject towards that one. This would result in an overall bias in the targets used, unless the pile were

regularly replaced. This would mean having extra piles of envelopes hidden somewhere.

2. The agent could guess which picture the subject would choose. This would be especially easy if he knew the subject well, or the subject had taken part in previous trials. He could then cause this picture to be selected by several methods e.g.

a. By marking the main pile of twenty envelopes and selecting the right one.

b. By taking an envelope from a drawer instead of from the main pile.

c. By concealing extra envelopes to use for the purpose (I thought of this some days later).

These methods would all be detectable. If (b) occurred a small envelope would disappear from the drawer *during* the trial instead of afterwards (during replacement). The pile might also remain at 20, instead of 19, during the trial. Or if it were reduced to 19, two envelopes would be used during one trial instead of one. Also the pile would become biased because the one removed would not match the one later replaced.

3. A most effective method would be for one person to arrange both to carry out a false randomization (as in 2) and also be present at the judging to 'help' the subject.

All these methods involve violations of protocol. Some would be easily detectable and I therefore decided to make certain simple observations which in no way interfered with the running of the experiment or with anyone's privacy. If I found no indications that any of them were happening, then I could be reasonably confident that the results were due to ESP.

The effects predicted were as follows:

A. The main pile might be marked.

B. The main pile might be replaced or partly replaced.

C. The main pile might be biased (this could arise from several methods). The only way to check this would be to open the envelopes which I did not wish to do (but see later).

D. There might be piles of extra envelopes around the room. I thought it improper to search for them and did not wish to do so.

E. Envelopes might disappear from the replacement drawers during, rather than after, a session.

F. Two envelopes, instead of only one, might disappear from the drawers for one session.

To check on these last two possibilities I decided to count the numbers of envelopes in each drawer both during each session and afterwards, and to watch the replacement procedure whenever possible.

To recap—my hypothesis was that if cheating were taking place I would expect envelopes to disappear from the drawers during, rather than after, a session, or for more than one envelope to be used for each session.

Results

The main pile did not seem to be marked and was not switched during these two days. I counted the envelopes in the drawers from session 8 onwards. The results are shown in Table 2.

Table 2. Numbers of Envelopes in The Main Pile and Drawers

Session	Main pile	Drawers A	B	C	D	Target
8	19	11	19	9	17	D
9	19	11	18	9	16	B
10	19	11	17	9	16	C
11	19	11	17	8	16	A

From this table it can be seen that between sessions 8 and 9 two envelopes, not one, disappeared from the drawers. One was a 'D' (which is correct to replace the 'D' which was target for session 8). The other was a 'B'. This was target for session 9, but of course the 'B' for that session should have come from the main pile; only being replaced later by one from the drawer.

I later observed the replacement procedure and this was carried out correctly. i.e. another 'B' was taken from the drawer and placed in the main pile.

If the 'B' for trial 9 had come from the drawers instead of from the main pile (as hypothesised in 2b above) this would probably result in there being an extra 'B', instead of some other letter, in the main pile. I was unable to check on this at this time.

I noted certain other problems all concerning the same trial. During this trial (No. 9) I stayed with the experimenter (K. R.) and watched the judging. Sargent was not officially taking part, but he came in during the judging. He said he wanted to help, because it was a particularly difficult session, the subject having said only a few words. He seemed to push the subject towards picture B. I wrote this observation in my notes at the time and K. R. independently mentioned it to me as well. Note that I wrote down this observation before I counted the envelopes in the drawer. Of course this ought not to matter because Sargent should not have had any way of knowing the identity of the target (but see later).

On the same trial there was also an arithmetical error: it was later discovered that the experimenter had added up the marks wrongly. Picture B had not been given the most marks and so this session should not have been a direct hit. When he discovered this Sargent checked the addition for all previous trials and found nothing else wrong. Rejudging would be one way to clarify whether there really was a good correspondence between the subject's mentation and the picture B. Sargent said that he intended to do this rejudging.

Observations Phase 3 Days 6–7

I intended to continue observing. I also considered asking Sargent whether we could open the envelopes in the main pile to see whether it had become biased as hypothesised. However, Sargent became ill with 'flu' and was away on Day 6. I was therefore unable to observe any more sessions or to ask him about the main pile.

In Sargent's absence I discussed the experimental design and its potential problems with Trevor Harley. I told him that I was worried that the main pile of

20 envelopes might become biased, and no-one would know it had happened.

He assured me that Sargent always did the replacement himself and that he would not make such errors. Nevertheless, he thought it was a good idea to open them to find out. He checked that there were new envelopes of the same kind available. I then opened all the envelopes. There were 19, the replacement for the previous trial not yet having been done. There should have been 4 'A's, 5 'B's, 5'C's and 5'D's. There were in fact 5 'A's, 6 'B's, 4 'C's and 4 'D's. As I had predicted there was an excess of 'B's.

Harley and I discussed the possible ways this error could have come about. These include:

1. Accidental errors made originally in the drawers.
2. Accidental errors made in replacement to the main pile. Two such errors could create the bias observed.
3. As a by-product of the methods (of cheating) outlined above.

We then opened the envelopes in the drawers. Drawers B–D were correct but the 'A' drawer contained 2 'D's in addition to several 'A's. Harley and I replaced all the letters in new envelopes and reconstituted the main pile correctly.

Because of finding these errors I discussed with Harley the reasons I had for worrying about them. I explained about the missing 'B' on session 9, and the other observations made concerning that session. Harley immediately recalled that on that session there had been a change from the official procedure.

Harley was to be agent. It was an experiment in which there was little time for the agent to do the randomization. Harley therefore asked Sargent to prepare things for him; apparently meaning him to get all the envelopes, tables and so on ready. In fact Sargent actually carried out the randomization and handed Harley the set of pictures and the small envelope. Harley took them and used them for that session. This should not have mattered since officially Sargent was to have no further role in that session. However, of course, we now knew that Sargent had come into the judging session on that occasion and had apparently 'pushed' the subject towards the correct picture.

The following day Sargent was still away ill. Harley and I wished to check up on some details of previous sessions and therefore looked for the book in which they were recorded. We could not find it, but in the process Harley found a sealed envelope, like those used in the randomization, under some papers. We decided to look for any further ones. We found a single one in a drawer and a pile of three under some papers. We opened them all. The single ones were a 'C' and a 'D'. The pile of three were all 'A's. We found no 'B's.

We discussed possible reasons for them being there. One possibility appeared to be the method 2c, outlined above. If there were no 'B's concealed, then only method 2b could be used and would result in a 'B' going missing from the drawers, as observed on trial 9. We discussed alternative explanations.

Harley said that the envelopes for this series of experiments had been specially prepared all at once and placed either in the main pile or the drawers. Envelopes of that size and colour had not been used in any previous experiment. He could think of no reasons for there being any extra ones around the room.

Two further sessions were conducted, by student experimenters, in Sargent's absence. These obtained ranks 3 and 4; both misses.

Explanations

When Sargent returned after his illness Harley presented him with the findings so far. These were:

1. The bias in the main pile and errors in the drawers.
2. The extra envelopes found around the room.
3. The series of events surrounding session 9.

Sargent denied that any of these errors had come about deliberately and supplied alternative explanations for them. I hoped that Sargent would write his own account and provide these explanations himself. Since he has never done so I shall try to be fair to what he told me. We now have two alternative hypotheses to account for the findings.

1. I had predicted that certain methods of cheating would lead to a bias in the main pile. I found that bias.

Sargent said that the errors in the pile must have come about by accidental errors in replacement.

He calculated the maximum size of any spurious effect that could be created by this bias and found it to be only 3 per cent; a negligible effect when the average hit rate was about 45 per cent. Clearly if the bias were accidental it could not account for the successful results. On the other hand if it came about as a by-product of those methods of cheating, a very large effect size could be obtained.

At this time the error in addition (mentioned above) was also found.

Neither Sargent nor I had any explanation for the 'D's in the 'A' drawer.

2. I had predicted that certain methods of cheating would necessitate having extra piles of envelopes hidden around the room. These were found.

Sargent explained that the extra envelopes had been left over from a previous experiment, although Harley had previously said that this was very unlikely.

3. It now appeared that on one session—number 9—the following events had taken place.

1. Sargent did the randomization when he should not have.
2. A 'B' went missing from the drawer during the session, instead of afterwards.
3. Sargent came into the judging and 'pushed' the subject towards 'B'.
4. An error of addition was made in favour of 'B' and 'B' was chosen.
5. 'B' was the target and the session a direct hit.

Sargent said he had done the randomization because Harley asked him to. Sargent said he had removed a 'B' because it was bent and therefore distinguishable from others. He said he had already told Harley about this. Harley now said he remembered being told although he had not remembered this previously when he and I discussed the problem.

Sargent said there was no harm in him coming into the judging since he did not know the identity of the target, even though he had done the randomization. He denied 'pushing' the subject.

There are therefore two hypotheses to consider. The hypothesis of cheating led to the discovery of the errors. It explains them fairly neatly and could, if extrapolated to the whole experiment, account for the large effects observed.

The alternative is *ad hoc,* and cannot account for the large effects (these would have to be attributed to psi). It would imply a good deal of carelessness in the running of the experiment.

I considered that the evidence was not conclusive in favour of either hypothesis and that more evidence was needed. I did not wish to make any accusation, or even implication, of cheating, without conclusive evidence that it had occurred. It therefore seemed essential to gain further information which might support one or other hypothesis, and in the meantime not to publicise the findings.

Further Hypotheses

There were several kinds of information which would be relevant:

1. Further observations of the experiments in progress. These were planned for a second visit of three weeks early in 1980. However, two weeks after I left Cambridge, Sargent informed me that he did not wish me to return, which of course I accepted.

2. The results of further experiments using the same procedure and subjects, but a different experimenter.

This was also part of our original plan, but did not take place for the same reason.

3. A full report by Sargent (and his colleagues) of their explanation of the errors.

In January 1980, I wrote a report for the SPR archives. This was to be available to SPR members on request, but I hoped it would soon be made redundant by a published version. Sargent and I agreed that we would each write our own version of the events. I wrote mine and sent it to him. He wrote an early (confidential) version, but never produced a final one. He continued to promise he would and therefore I waited and did not publish my own account.

When it became clear that Sargent was unlikely to produce a report, I discussed with Harley the possibility of publishing a joint account. We differed in some respects but agreed that we could write a report together if the points of disagreement were made clear. Harley did not write a report. I finally concluded that no written explanation was likely to be forthcoming from either Sargent or Harley.

4. Further analyses of raw data from previous experiments.

There were several ways in which the raw data might help to test the hypotheses. For example, according to some methods of cheating one would expect the most popular picture in any set to have been target more often than predicted by chance. I asked whether I could check this. However Harley said that the pictures in each set were changed from time to time, without any record being kept, and that it would be impossible to check this from the existing records.

Another hypothesis was that, if one person were cheating and pushing the subject towards the target, rejudging should give poorer results than the original ones. This would be easy enough to do and Sargent said that he intended to do it. However he never published the results of any rejudging.

Thirdly, if one person were cheating, the most significant results should occur when they were acting as agent or experimenter, though of course this could also occur because of a psi-mediated experimenter effect. In fact there is evidence that scores were higher when Sargent took part both in the few sessions observed during this visit and in published data (Ashton, Dear, Harley and Sargent 1981). I hoped to be able to check the entire data base for this effect. This would mean having the Blue data book in which the names of all participants are recorded.

Finally, another suggestion was made by Parker and Wiklund (1982). Cheating could take place by manipulation of the randomisation combined with knowledge of the subject's likely responses (as in 2a-c above). The easiest way to find this out is by looking at the subjects' responses on previous trials. Wiklund and Parker suggested that in those trials where Sargent was responsible for the randomisation, and the subjects did not make direct hits, there would be above chance scoring if the target were matched with the subject's mentation on a previous trial (Parker and Wiklund). This could be checked from the raw data and they therefore asked Sargent for those data.

These suggestions provide definite ways in which the implications of cheating could be lifted. If Sargent supplied the raw data other researchers could check them for these effects. If these effects were found, that hypothesis would be strengthened. If they were not found then the cheating hypothesis would lose much of its force.

I kept hoping that this would happen and the truth become clearer. However Sargent refused to make his data available. Several informal requests for the data were made. Then when these failed to elicit any data, official requests were made through the Parapsychological Association. Sargent still did not supply the data, nor any reason for withholding them.

In 1984 the PA Council asked Martin Johnson to head a committee to investigate the case. The final report of this committee is now available. Council reprimanded Sargent for failing to respond to their request for information within a reasonable time.

In view of this lack of cooperation it is not possible to test any of these hypotheses against the data. Also there now seems little hope of obtaining any new evidence and therefore we must assess the case on the basis of what evidence we already have.

I have been criticised for not publishing a full account earlier. I hope I have now made clear my reasons. I did not wish to publish something which discussed the hypothesis of cheating, (a) while there were still promises that others would supply alternative explanations for my findings and (b) while there was still some hope that further evidence would come to light.

I think there is still doubt as to the correct hypothesis. However, any hope that this will be speedily resolved now seems to be unrealistic. I am therefore presenting the evidence I have, as accurately as possible. I hope that others will add their versions to mine.

Implications

There has recently been considerable controversy concerning the value of the ganzfeld database in providing evidence for psi. The many experiments involving Sargent as experimenter form a very substantial and important proportion of that database. According to Hyman (1985) Sargent's 9 studies and Honorton's 5 account for one third of the total. According to Honorton (1985)

Sargent's experiments have the second highest effect size, after Honorton's own. If Sargent's findings were removed from this database it would be considerably weakened as evidence for psi.

Brain and Perception Laboratory
University of Bristol, Bristol.

REFERENCES

Ashton, H. T., Dear, P. R., Harley, T. A. and Sargent, C. L. (1981). A four-subject study of psi in the ganzfeld. *JSPR,* **51**, 12–21.

Blackmore, S. J. (1980). *Extrasensory Perception as a Cognitive Process.* Unpublished PhD Thesis, University of Surrey.

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

Hyman, R. (1985). The Ganzfeld Psi Experiment: A Critical Appraisal. *JP*, **49**, 3–49.

Parker, A. and Wiklund, N. (1982). The ganzfeld: A methodological evaluation of the claims for a repeatable experiment. Unpublished.

Sargent, C. L. (1980). *Exploring Psi in the Ganzfeld.* Parapsychological Monographs No. 17.

Sargent, C. L., Harley, T. A., Lane, J. and Radcliffe, K. (1981). Ganzfeld psi optimization in relation to session duration. *RIP 1980*, 82–84.

Sargent, C. L. and Matthews, G. (1982). Ganzfeld GESP performance in variable duration testing. *RIP 1981,* 159–160.

CHEATING, PSI, AND THE APPLIANCE OF SCIENCE: A REPLY TO BLACKMORE*

by TREVOR HARLEY AND GERALD MATTHEWS

ABSTRACT

Blackmore describes her visit to the Cambridge laboratory in 1979 and discusses evidence for what she calls a 'cheating hypothesis'. This is that certain anomalies which she discovered are best accounted for in terms of experimenter cheating. We demonstrate that the so-called 'cheating hypothesis' is not a hypothesis in the traditional scientific sense of the word, and that she is guilty of extreme prejudice in her reporting of the events and in their interpretation. We then analyze some data which refute her claims empirically. The best interpretation of events is also the most obvious—minor experimental error.

Blackmore (this issue) reports her visit to Cambridge in November 1979, and concludes from this that not only must the data and interpretations of all the ganzfeld experiments carried out there be discarded, but also implies that there is a strong possibility of fraud on the part of Carl Sargent. As we are referred to in her report as co-experimenters of Sargent, such an accusation questions at least our competence. This reply demonstrates that the only sensible interpretation of Blackmore's observations is a 'random errors' hypothesis. Furthermore, Blackmore is prejudicial in her reporting of the phenomena, and altogether too sanguine in her interpretation of them.

Our reply is organized as follows. First, we will discuss some general issues concerning Blackmore's observation and her interpretation. Second, we will examine specific points. Third, we will analyze some previously unpublished summary data which show that there is no empirical foundation for her claims. Finally, we will mention the wider implications of her approach.

GENERAL ISSUES

Blackmore's paper can be summarized briefly thus. She visited the Cambridge laboratory in 1979 with the stated intention of comparing it with her own laboratory, with a public view to establishing possible reasons for the difference in the results of the two research groups—the Cambridge laboratory very successful, her own very unsuccessful. She observed a number of irregularities from the stated protocol, and interprets these irregularities as demonstrating support for the cheating hypothesis. We will examine these irregularities in detail below, but they centre around the randomization technique used in general, and one session in particular.

The first point to be noted is that Blackmore's use of the phrase 'cheating hypothesis' is far too casual. It is not a hypothesis in the normal scientific sense of the term, since it is unclear what evidence would be sufficient to falsify it. The

* Acknowledgements
The authors gratefully acknowledge financial assistance from the SPR. The order of the authors was randomly determined. Reprint requests and other communications to: Dr. Trevor Harley, Department of Psychology, University of Warwick, Coventry, West Midlands CV4 7AL, U.K.

'cheating hypothesis' really comprises multiple hypotheses, some contradictory, which taken together generate so many predictions that blanket coverage of the phenomena is obtained. Since the whole point of the experimental protocol is to prevent sensory leakage, almost any protocol violation can be taken, post hoc, as supporting the cheating hypothesis. Conversely, had Blackmore in fact not observed any protocol violations, the cheating hypothesis could easily have been rescued. Sargent could have used a technique undetectable by Blackmore's methods (we provide an example of this below), or taken other precautions to avoid detection. Thus the function of Blackmore's 'cheating hypothesis' is simply to justify the post hoc interpretation of protocol irregularities as cheating, while labelling an interpretation of the same irregularities in terms of random error or carelessness as arbitrary and ad hoc. If her 'predictions' were truly generated in advance, one may wonder why they were not lodged with some independent body, such as the SPR. Any resemblance here to the scientific method is entirely coincidental.

If one were to try and rigorously construct a 'cheating hypothesis', it is likely that one would want a single hypothesis stated somewhat as follows. The experimenter will in a clandestine fashion manipulate some specified portion of the experimental design so that his or her experimental predictions will appear to be verified even if that does not in fact occur, while taking trouble to conceal his or her manipulations from others. Hence, if we were to observe the experimenter closely we would expect to find a consistent pattern of behaviours. The 'carelessness hypothesis', on the other hand, predicts an essentially random pattern of anomalies. This is in fact what was observed by Blackmore, as we will demonstrate in detail in the next section. her use of the word 'prediction' in this context is far more rhetorical than scientific.

It should be noted that the occurrence of random errors in such experiments is unfortunate, but perhaps not surprising. As Blackmore notes, there were elaborate precautions built into the design to preclude cheating and sensory cueing. A side-effect of this level of elaborateness is complexity, and increased complexity usually leads to an increased likelihood of errors. It should also be noted that most of the sessions monitored by Blackmore were from experiments which were part of undergraduate student projects. Anyone involved in University research will testify to the dangers of leaving experiments to students. As Blackmore herself concludes, for her method 1 at least, there is no way in which the observed anomalies could have given rise to the observed size of effects in these experiments if they were genuine errors. She neglects to say, however, how the same anomalies could have given rise to statistical effects of this magnitude if these anomalies were due to cheating.

Second, Blackmore completely overlooks the scientific context within which these experiments occurred. The experiments under discussion were not designed merely to demonstrate the existence of psi, but to develop a theory which specifies the conditions under which psi is expected to occur. Hence hits were not predicted on every occasion. Take as an example Blackmore's session 9 in the KR experiment (subsequently published as Sargent, Harley, Lane, and Radcliffe, 1981), of which Blackmore makes so much. First, the subject, RC, was an introvert, a strong contra-indicator of hitting in a session, and a novel finding of which Sargent was rather proud. Also from Blackmore's own records it can be seen that this particular trial was a 15 minute session, a condition in which subjects were *predicted* (in the correct use of the term) *not* to display psi, because at this stage the altered state would not be sufficiently developed. Hence *if* Sargent were to cheat and attempt to improve the hitting rate, he certainly would not do it in the session which is the centre-piece of Blackmore's presentation.

Blackmore's analysis of her own data is prejudiced throughout. Furthermore, there are discrepancies with a report compiled by one of us (TH), which was also compiled immediately after her visit to Cambridge. In particular, Blackmore barely mentions the fact that Sargent told TH that he had replaced a B envelope, because it was marked, well *before* this session took place.[1] It is illuminating to note that the reason he did this substitution was because the original was distinguishable from another envelope by virtue of the corner being bent. It was another prediction of Blackmore's 'cheating hypothesis' that cheating could have occurred by use of marked envelopes. Hence if Sargent had not made this switch, she would have found him guilty of fraud by use of marked envelopes. As he did make the switch, she instead favours cheating by manipulating the target pile! In any case, Blackmore dismisses Sargent's explanation, claiming that she cannot see why a bent and therefore distinguishable card should be removed from the target pile. This is quite extraordinary given that elsewhere another of her cheating hypotheses (2a) was that the main target pile would contain marked cards. This is yet another example of poorly formulated and over-generalized hypothesis formation.

The beauty and power of her 'cheating hypothesis' is here evident. The rest of us must often wish that the hypotheses with which we work were so difficult to falsify.

It is possible that one reason for this lack of charity in her interpretation is that Blackmore was influenced by Sargent's later lack of co-operation. Whereas this is understandable, it should not have affected her reporting. It must be stressed that subsequent events are totally irrelevant to the interpretation of Blackmore's observations.

Specific Issues

We can now examine in more detail some of the particular points made by Blackmore in her report.

(1) *The mystery of the missing B*

Blackmore is somewhat disingenuous in her reporting of this. First, we have mentioned above that she barely addresses the fact that one of us (TH) was told by Sargent that he had replaced a B envelope *before* the session. Second, the given explanation is perfectly reasonable. Third, Blackmore also states (3.2) that the B went missing *during* session 9. This is quite wrong, as surely all Blackmore could have known was that the B went missing after she last counted the piles. Indeed, earlier in blatant contradiction of this, she says that two envelopes disappeared from the drawers *between sessions 8 and 9*. From her table 2, it can be seen that she only counted the piles once per trial, yet earlier in the text she states that she decided to count the numbers of envelopes both during the session *and*

[1] It is likely, in fact, that TH *was* actually present when Sargent made this replacement.

afterwards. Despite these gross inconsistencies, we can infer that she only checked the drawers in between sessions. In her text this inconsistency is resolved in favour of the more damaging version: yet another example of her grossly prejudicial reporting.

(2) *The mystery of the main (target) pile distribution*

This appears to be partially consistent with the main hypothesis. However, Blackmore can provide no account of why there should also be an extra A in this pile. This anomaly is therefore better explained by the 'random errors' hypothesis. The previous 3 trials with CLS as randomizer used D, D, and C, as can be seen from our appendix A. (The three trials observed by Blackmore all came from this single experiment, reported as Sargent and Matthews, 1981).

Blackmore fails to state that this distribution *rules out* her 'method 1' of cheating entirely. This is another example of selective reporting.

(3) *The mystery of the extra envelopes*

These in no way constituted a violation of protocol. They could easily be from other experiments, and Sargent's explanation is perfectly reasonable. (Sargent was also involved in piloting a number of other experiments involving neither TH nor GM). Indeed, if cheating were occurring by 2b (taking envelopes from the drawer), it is implausible that it should simultaneously occur by 2c (concealing extra envelopes), as implied by Blackmöre. Blackmore also admits that 2c was thought of 'some days later', so this can hardly be thought of as a prediction, even with Blackmore's strange use of the word.

For future experimenters, here is a much better way of cheating than any of the methods proposed by Blackmore. The probability of drawing the desired card from the target pile, without replacement, in 8 attempts, is 0·949. All that is necessary is that the randomizer has a small supply of envelopes concealed upon their person, which can then be used to replace the wrong cards. This obviates the need for the complications of Blackmore's mechanisms, and is much safer, being detectable only by a body search.

(4) *The mystery of session 9*

The main point here is that Sargent would *not* have predicted a hit in this condition. As we have said, this was an introvert subject in a short-trial condition.

The alleged pushing was unsuccessful, and the B was still put second, even allowing for the arithmetical error. Blackmore states that 'the experimenter', presumably KR, was responsible for this addition at the end of the experiment. Are we then to believe that KR was also involved in cheating? Yet KR also seemed to be worried that Sargent was 'pushing' the subject towards the B! Surely arithmetical juggling of scores must be the clumsiest possible method of ensuring success.

It is straight-forward to account for the alleged 'pushing'. The pictures used were complex, and often contained a large number of elements. Naive subjects in particular tend to develop a cognitive set such that they preferentially attend to certain elements, while ignoring others. The role of the experimenter, then, is simply to point out possible correspondences between the transcript and the elements of the picture which the subject seems to be ignoring. No pressure is applied on the subject to accept these correspondences. It should be obvious, therefore, that if particular subjects (including relatively experienced ones) have a strong but irrational preference for a particular picture, merely pointing out correspondences with another picture might appear like pushing them towards that other picture especially to an inexperienced observer. It should not be forgotten that KR was an inexperienced experimenter, and that as this was an undergraduate project contributing to KR's final degree, Sargent had a moral responsibility to ensure that KR was acting both correctly and maximizing the chance of success.

Finally, Blackmore fails to report that when this session was independently rejudged, the new judge, DG, rated the session a stronger hit than those participating in the original session.

(5) *The mystery of the best way of cheating*

Throughout it can be seen that Blackmore does not really have any clear idea of how cheating might be occurring. She provides a mish-mash of hypotheses, which we have argued are much better explained in terms of random errors. Neither does she point out that if one were to cheat, one would hardly do so in front of a known sceptic who has come with the public purpose of closely monitoring procedures. Furthermore, she mentions but does not discuss the conclusion that the observed bias in the distribution could only produce a bias of 3 per cent, if her method 1 were used, which as Blackmore rightly observes, is negligible, and could not account for the average hit rate of 45 per cent. Indeed, we are not clear exactly how cheating was supposed to be occurring according to Blackmore. She seems to make something of each of the skewed distribution, the missing envelope, the supposed pushing, and an arithmetical error. It is surely preposterous to propose with any seriousness that anyone could be so stupid as to have to resort to all these methods of manipulation when much simpler, more reliable, and less detectable methods can be used.

A Reanalysis of some Summary Data

Appendix A shows the complete record of one of the experiments observed by Blackmore (her experiment 1, published as Sargent and Matthews, 1981). An analysis of these data clarifies and extends much of the above.

The magnitude of the effects necessary to achieve particular goals is consistently ignored in Blackmore's report. It is doubtful how effective either pushing or guessing a subject's preference for a picture can be. Blackmore presents no psychological data.

If the probability of succeeding in this tactic of guessing subjects' preferences is assigned the very generous estimate of 0·5, then intervention will be necessary on 18 trials to get the 11 hits necessary for overall statistical significance out of 26 trials. One might certainly want a stronger effect than this, and then the number of trials requiring intervention will increase. Hence intervention on a large scale will be necessary.

In the experiment shown in Appendix A, Sargent was present on 22 trials, 17 of them as experimenter. This was not surprising, as GM would naturally have acted as agent for most of the 13 subjects who were his friends and acquaintances. This would have left Sargent only 5 trials when he was agent when he could have intervened. This is insufficient, given a MCE (mean chance expectance) of (number of trials)/4 for the experiment, as the expected 5·25 hits in the remaining 21 trials would still leave the experiment short of the target 11 hits. Thus, even if he were capable of a perfect guessing of the subject's choice, this method would be doomed to failure. Furthermore, he was accompanied by RH, acting as agent's experimenter, in trial 16, which presumably reduces the number of trials in which he could have intervened by this method to a mere 4.

Table 1 shows how the success of Sargent as agent compares with that of other agents.

Table 1. Comparison of hit rates for different agents

Agent	*Hits*	*Misses*
CLS	3	2
Others	9	12

Sargent's hit rate is 3 hits out of 5 sessions (exact binomial probability ·0879), while other agents are in fact closer to significance (9 out of 21; exact binomial probability ·0561). A Fisher exact probability for the differences in this distribution is not significant. Neither is there a difference in the mean z-scores between Sargent as agent, and other agents, for these sessions (CLS = + 1·08; others = + 0·38; t = 1·51, 24 df, p = ·144, 2-tailed).[2]

This implies that if Sargent were cheating, he could not be doing it in the role of agent alone, and hence the idea that he was not following protocol in target selection must be rejected. Hence if he were cheating, he must have either been cheating as both experimenter and agent, or as experimenter alone. The emphasis therefore falls upon cheating as experimenter, and for this, given all the above, he could only rely upon a large bias in the main (target) envelope pile, and an almost precognitive ability to judge subjects' preferred pictures. In this experiment, Sargent was experimenter 17 times. Assuming chance performance on the other 9 trials, Sargent would have had to have guessed subjects' preferences with the incredible probability of 0·51 to obtain the 11 hits necessary for statistical significance. If he were also cheating with a 60 per cent hit rate in the trials in which he was agent, this probability only falls to 0·41. Hence we must look elsewhere for the alleged manipulation.

In addition, if the pile were biassed throughout the experiment, one would certainly predict one target to be more common than the others. Out of 26 trials, the distribution for all 26 trials was as shown in Table 2.

Table 2. Distribution of targets throughout an experimental series

A	B	C	D
7	7	5	7

[2] There is thus no difference, either in the proportion of hits or in the significance of the score between the trials where Sargent was acting as agent versus the others.

This distribution is clearly random (chi-squared = 0·46, 3 df, $p > 0{\cdot}2$). Hence it would have to be argued that the pile would have to be periodically changed, so that the biassing letter would change every few trials or so. (Note that if the criterion of bias is simply picture popularity, this would also involve rather implausible manipulation of the target sets so that there are 'sets of sets' where the biassed letter corresponds to the most popular picture. Hence a method of manipulating target set selection must also be proposed). But if this were the case one should clearly expect non-random sequences.[3] If the experiment is divided into quarters, the distribution is shown in Table 3.

Table 3. Distribution of targets through four quarters of an experimental series

	1st.	*2nd.*	*3rd.*	*4th.*
A	3	2	2	0
B	1	1	2	3
C	1	1	1·5	1·5
D	1·5	2·5	1	2

This also is clearly random. The largest run throughout the experiment consists of three As, which is not significant by a Runs test. Neither is there any effect of A's and B's versus C's and D's, nor of A's and C's versus B's and D's, also by the Runs test.

It should also be noted that during Blackmore's visit, trials 8 through 13 of the experiment summarized in the appendix were performed. Prior to Blackmore's session 9, the targets were in sequence C, D, A, and D, with no B.

The only possibility remaining is the technically very difficult one of switching the pile from session to session. Fortunately Blackmore states that she checked for this, and that it did not occur.

Conclusions

Blackmore's report is loaded with prejudicial reporting and inconsistencies. If analyzed correctly, the data clearly show that her observations are best explained by a 'random errors' hypothesis. As Blackmore herself states, these can in no way account for the magnitude of effect demonstrated in the Cambridge experiments.

Taking Blackmore's data and the appendix together all of Blackmore's specific 'hypotheses' are refutable. It is, of course, impossible to rule out the 'cheating hypothesis' altogether, as it is presumably always possible to think of more and more contrived explanations. Her hypotheses 1 to 3 are ruled out as follows:

Hypothesis 1: Blackmore's data show only a trivial bias in the distribution of target letters at one given time. The data in the appendix show no significant bias in the distribution of letters across a single experiment or portion of an experiment.

[3] Blackmore entirely fails to address the problem of target *set* selection.

Hypothesis 2a: This was checked for and not found by Blackmore.

Hypothesis 2b: Blackmore failed to check for the disappearance of an envelope during a session. The exact distribution of letters is best explained by random error.

Hypothesis 2c: The presence of extra envelopes does not require explanation and was in no way a violation of the protocol. In any case, if one were using 2b, why would one also need 2c?

Hypothesis 2, all versions: They are all also disconfirmed by an analysis of the data in appendix A, which show that the trials in which CLS acted as an agent could only have added 1·75 in excess of MCE (Mean Chance Expectance).

Hypothesis 3: The only evidence offered for this by Blackmore are the events of session 9. Sargent was randomizer during 4 other trials observed by Blackmore. In 3 of these her table 1 shows that she would have observed any further violation of protocol. As above, it is inconceivable that an experimenter would fraudulently generate results in an opposite direction to that predicted.

What can be learned from all of this? Much has been said about this case, and this is no forum for its repetition. (We refer the interested reader to the Parapsychological Association). One may criticize the randomization technique as it allowed minor errors to occur, but any such technique which is complex enough to avoid worse problems and which at some stage involves human intervention, is bound to allow scope for errors. We feel that there is room for constructive criticism of the ganzfeld, such as Hyman's (1982) critique. Indeed, we would go along with many of those criticisms. We would particularly welcome suggestions for the improvement of randomization techniques, which we also perceive as a methodological difficulty with the ganzfeld. However, the destructive and negative criticisms which seem to dominate parapsychology are perhaps a sign that it lacks proper scientific maturity. Primarily, our overall reaction is one of surprise: surprise that observations which are so clearly accounted for by random errors should be interpreted in so hostile and negative a fashion. Until parapsychologists can avoid such meaningless debates, the subject has no future.

Trevor Harley
Department of Psychology
University of Warwick
Coventry CV4 7AL

Gerald Matthews
Division of Applied Psychology
University of Aston
Birmingham B4 7ET

REFERENCES

Hyman, R. 'Does the ganzfeld experiment answer the critics' objections?' Proceedings of the Society for Psychical Research and Parapsychological Association Conference (Volume 1), Cambridge, 1982.

Sargent, C. L., Harley, T. A., Lane, J. and Radcliffe, K. 'Ganzfeld and psi-optimization in relation to session duration'. *RIP 1980.* (1981).

Sargent, C. L. and Matthews, G. 'Ganzfeld performance in variable duration testing'. Paper presented at the Parapsychological Association Meeting, Syracuse, New York, 1981.

APPENDIX A

Sess	S	Es	Ea	Date	Ratings A	B	C	D	Target	Rank	Z-score
1	GM	CS	TH	4·11	35	75	70	25	C	2	+0·87
2	CS	GM	TH	8·11	90	50	08	37	A	1	+1·48
3	SK	CS	GM	9·11	60	69	58	85	A	3	−0·75
4	TH	GM	CS	11·11	60	78	40	45	A	2	+0·29
5	AS	CS	GM	13·11	18	20	15	75	D	1	1·73
6	CB	CS	GM	15·11	10	70	30	87	B	2	+0·68
7	PG	CS	GM	16·11	70	30	05	95	D	1	+1·29
8	AT	GM	TH	22·11	50	25	65	35	C	1	+1·40
9	MN	GM	CS	23·11	05	10	50	75	D	1	+1·38
10	JB	CS	GM	24·11	55	80	79	32	B	1	+0·94
11	JF	CS	GM	25·11	75	20	50	70	A	1	+0·98
12	HA	GM	CS	25·11	86	24	22	68	D	2	+0·65
13	SB	GM	TH	29·11	45	35	60	55	A	3	−0·39
14	JL	CS	GM	1·12	35	75	55	32	A	3	−0·82
15	RH	CS	GM[1]	10·12	25	45	70	20	B	2	+0·25
16	SB	GM	CS[2]	10·12	75	60	40	30	A	1	+1·36
17	JA	CS	GM[2]	13·12	30	00	50	60	B	4	−1·31
18	IH	CS	GM	16·12	30	90	50	75	D	2	+0·60
19	LS	CS	GM	20·12	55	40	80	50	C	1	+1·40
20	CC	CS	GM	9·1	60	70	20	35	C	4	−1·33
21	KR	CS	GM	19·1	55	30	40	60	D	1	+1·15
22	GS	TH	GM	24·1	04	70	40	10	C	3	−0·01
23	MS	CS	GM	26·1	60	40	80	25	B	3	−0·54
24	DG	GM	CS	26·1	10	90	15	20	B	1	+1·72
25	ES	CS	GM	27·1	48	15	45	60	D	1	+1·09
26	DJ	CS	GM	31·1	03	05	10	25	B	3	−0·67

Notes to Appendix A

1. Dates refer to 1979–19^0.
2. Abbreviations used: TH = Harley, GM = Matthews, CLS = Sargent, Sess = Session number, S = Subject, Es = Subject's experimenter, Ea = Agent's experimenter.
3. The data from this table were later published as Sargent and Matthews, 1981.
4. On trial 16 Sargent was accompanied by RH and therefore would in any case be unable to cheat.
5. Explanation of footnotes:

[1] Accompanied by SB = Blackmore

[2] Accompanied by RH

SCEPTICAL FAIRYTALES FROM BRISTOL*

by CARL SARGENT

INTRODUCTION

This is a rejoinder to Blackmore's (1987) paper. As a preliminary to dealing with the details of this paper, there are two important considerations to be kept in mind by the reader. First, there are two versions of Blackmore's paper, a 1979 report for the SPR (originally confidential but circulated by Blackmore in later years) and the 1987 version, which I believe to be the same as that written in 1985. The differences between the two are of some importance; the more sceptical tone of the 1987 version is extremely marked and in places this affects the content also (see section on Blackmore's suppression of evidence below). Second, and this is crucial, *Blackmore has nothing save for her own testimony* to recount to the reader. In my account, I shall be able to draw on diverse supporting testimony from others and other sources to support my claims in several instances. I too have contemporaneous notes to draw on but I alone also have more than this—Blackmore does not. Since, as I shall show, Blackmore's reliability and integrity are suspect, this is an issue of major importance.

I shall first deal with errors of commission and omission in Blackmore's account. I shall assume the reader is familiar with our experimental procedures from Sargent (1980).

SIMPLE ERRORS IN THE TALE

Blackmore's account is littered with errors of commission and to save the reader the boredom of going through all of them I shall just note six, of diverse importance, here although others will come to light later. Again, keep in mind that Blackmore only has her own testimony; my corrections of her errors have a wider data base.

To start with two very simple examples, Blackmore makes errors in describing our judging procedure. She claims that ratings were made on a 1–100 scale; actually it was 0–99, which is stated quite explicitly in the relevant experimental reports. A second error lies in her account of correcting an error in an experimental procedure where the judging set of pictures was found to be one picture short; she claims that the agent was told to take a duplicate (if he had one) and 'place it on the ground floor'. That this must be wrong is obvious, since Blackmore specifies no location on that floor; and indeed the agent was asked to leave the picture in the seminar room on the *first* floor (the location of this can be ascertained by a floor plan if anyone is pedantic enough to bother). There are other problems with this account, as we shall shortly see. She claims that her strategies for exploring fraud were 'simple and unobstrusive', a claim which Trevor Harley and I read with incredulity; if thoroughly searching someone's office and making covert checks as 'simple and unobtrusive' then the moon is made of marmalade. Another error comes not from this 1987 paper but from Blackmore (1986)'s account of her visit, where Trevor Harley and Gerry

* This title is lifted from Chapter 3 of the section 'Conversations with Illiterates' from Paul Feyerabend's *Science in a Free Society* (NLB, 1978), with apologies. It seemed doubly appropriate.

Matthews inform me that Blackmore could not even describe their physical appearances properly, making errors over such simple matters as hair colour and length. This is a relevant citation here because we are in the business of establishing how unreliable Blackmore's testimony is. Finally (at this stage) we can return to the session in which it was necessary to telephone the agent to see if he had both copies of a picture missing from the judging set of pictures; in addition to the simple error noted above, Blackmore makes others. These are of major importance. First, the agent was *not* asked to see whether he had a duplicate picture. After all the experimenter did not need to know this! All he needed was a copy of the picture missing from the judging set. Indeed he should not have asked (and didn't) about duplicates, *for the agent wouldn't know unless he had opened the appropriate envelope* (which would mean that this picture was the target). So, asking about a duplicate could have meant asking the nature of the target, and the experimenter did not do this. Blackmore does not inform the reader that the agent was also told as soon as he lifted the 'phone not to speak (agents didn't do this, but the call was obviously extremely early since judging hadn't even started, and we wanted to be sure that the agent didn't make any exclamation through surprise). Neither does Blackmore tell the reader that the agent was told to coat the picture concerned in fingerprints, so as to be sure to eliminate any possible handling cues.

Just looking at this one session alone. Blackmore makes no fewer than two errors of commission and two of omission. Her account clashes with my testimony, that of the agent, and the floor plans of the Psychology building. This is quite a fair start for a fairytale. We shall find other mistakes later, but I now wish to outline three errors where Blackmore's inability to give a correct account has arguably more sinister implications; we now deal with three instances of suppressed evidence.

Blackmore's Suppression of Evidence

1. *Rejudging:* This is a highly important example. This concerns session 9, the one in which Blackmore claims to have observed experimenter bias (I shall deal with this more fully later). In her 1979 report, Blackmore details for the reader an empirical check she made on this bias. Keep in mind that she claims that I 'pushed' the subject into placing the correct picture first. One way of checking on this would, of course, be to have the session independently judged. *This is exactly what Blackmore did.* She had a Ph.D. student in the lab., who had taken part in several experimental sessions in all roles, to rejudge the session. *He placed the correct picture first in his judging.* This, of course, is an independent empirical observation which questions the validity of Blackmore's alleged perception of bias. In the 1979 report, Blackmore cites this procedure. *In the 1987 paper she has suppressed it completely.* There is no mention of this important observation to be found in the 1987 report. Why not?

Now, to be fair, Blackmore bungled the original rejudging. Or so she claims; she asserts in her 1979 paper that the Ph.D. student could not remember afterwards whether he had not heard any mention of the session in question (even though the time lag was very short), so that it cannot be a certainty that the judging was truly 'blind'. So, Blackmore may have decided to delete her earlier account because the procedure wasn't fully correct. There are two rejoinders to this defence (if it is used). The first is that this observation is still an important one and should have been cited with the concern over blindness as it was in the 1979 report. That it has not been so reported raises the obvious question for the reader of either report, 'what else did Blackmore do that she isn't telling us about?' Quite a lot, actually, but even I cannot be certain of all of her unreported observations and actions. Second, note that Blackmore has played a classic sceptical 'heads I win, tails I can't lose' trick on us here by asking the judge about possible leakage *after* the judging. In this way, his placing it first can be discounted. One has to wonder whether she'd have asked about leakage if he had placed it second, I shall content myself with noting that a competent approach would have been to ask about possible leakage *before* getting rejudging done—this is an obvious part of method, after all. This suppression of relevant evidence is worrying. There are two worse cases to follow.

2. *Session 7:* Blackmore gives no account of this in either of her reports although she has verbally accepted my account on two occasions. In this session, I was the experimenter. I was concerned that the subject was very much biased against one of the pictures in the judging set because it was a rather dull black-and-white postcard showing a Cambridge scene which the subject made quite clear his dislike for, I considered that, as a naive subject, he was allowing his dislike to overcome accurate judging of correspondences between his mentation and the picture. In my view I did go overboard in trying to compensate for this, and we have some important independent evidence. Blackmore herself was present at the judging, and she was sufficiently concerned over my activity in the judging to ask me straight out which picture *I* thought was the target. This is important; it is the only time when we have any expression of concern at the time on her part over this issue (she made no such query for session 9, even though we were both present and she claims to have perceived similar bias again). As it transpired *I was wrong* and the subject chose the correct target and rate and ranked it first. This is a crucial observation, for it means that Blackmore has allegedly perceived judging bias on my part twice. In both cases the bias concerned whether a picture should have been placed first or second. Against a 50 per cent chance hit rate of my alleged bias pushing the picture in the correct direction we find an empirical 50 per cent rate. Suddenly Blackmore's allegations of judging bias look pretty frail. Doubtless this is why no mention of session 7 appears in either of her reports. (After all, she claims to have comprehensive notes about the sessions she witnessed.)

3. *Data Checking:* Blackmore makes much of the fact that I won't release data for re-analysis, and one of the reasons (again I will deal with this in detail later) needs discussion here. *Simply, Blackmore spent hours looking at the data when she was in Cambridge.* She did this on Day 7 in the company of my colleague Trevor Harley. She spent an afternoon going through the data in the master record book, making notes and performing statistical analyses (unfortunately we do not know what these were). Obviously she must have had *some* idea of what she was doing and what she was looking for. Equally obviously she didn't find it. Blackmore has verbally defended her failure to note this activity in her 1979 report on the incredible grounds (well, I find them incredible anyway) that she kept no records of what she did with the data. First, this would not be an acceptable defence even

if it were true; this should have been mentioned in her report given that she makes much of a need to do precisely what she did in the first place. Second, considering that she claims to have detailed notes about everything else, one's reaction to the claim that she has no adequate notes about this part of her enquiries can only be that of Senator Daniel Moynahan faced with Reagan's account of Irangate: 'it's unbelievable and frankly I don't believe it'.

My casting doubts on this score may not be accepted by the reader who may make a different judgement. I accept this. But it is a matter of fact that Blackmore carried out analyses of our data and that these were unreported in her accounts of her visit, and they must have had relevance to what she was doing. Their nonreporting seems to me completely indefensible. Again, note that it is not my testimony which is involved here but that of my colleague Trevor Harley who was present when Blackmore had the data. Lastly, the fact that Blackmore has scrutinized our data has obvious implications for re-analysis which I shall spell out later.

I shall now deal with Blackmore's account of session 9 before dealing with some further empirical issues and then addressing the credibility of the 'fraud hypothesis' as Blackmore terms it.

Session 9

Blackmore notes (correctly) that I undertook the randomization for this session. Trevor Harley had indeed asked me to prepare things for him, and I took this to mean having everything ready for him to leave for the other building where he would act as agent in the session. Blackmore claims that Trevor 'apparently' just meant me to get out the envelopes and tables. Actually, neither Trevor nor I can recall exactly what he *did* mean, but if he did mean what Blackmore claims he did this was a simple misunderstanding. There are important aspects to this which Blackmore does not mention.

First, why Trevor Harley should have asked me to help at all. In the experiment in question, 50 per cent of the sessions were 15 minutes long and the agent had to do the randomization and get to a room in another building within six minutes of the session starting. This was quite a rush and since I was sitting working in my office it is not surprising that Trevor Harley should have asked me to help. Second, and this is quite crucial, he asked me immediately before the session and this was not a regular procedure (although we were beginning to find the schedule very tight). The necessary implication of this, of course, is that it tells very strongly against any premeditated fraud! By this time the bent B-envelope in the deck *had already been destroyed*; Trevor Harley had observed my doing so (see below on this). Thus, linking the destruction of this faulty envelope with my conducting the randomization is absurd if Blackmore tries to construct a generalized argument out of this (in any case her attempts to do this are completely incoherent as we shall see when we examine the 'fraud hypothesis'). Lastly, although Blackmore doesn't mention this in her text, tucked away in her Table 1 it may be seen that I also acted as randomizer for session 1 when the same agent was involved; in this instance I did so simply to illustrate the procedure to Blackmore, who had just arrived. In this session I most certainly did *not* enter the judging room until the session (and judging) were completed; I joined Trevor Harley and Blackmore when they returned from the other building. But this was also a direct hit and unlike session 9 a correctly judged direct hit. This needs to be pointed out to show that the link, randomizer-present at judging-hit scored, is not necessary even in Blackmore's own recording for the link, randomizer-hit scored, to exist (which obviously detracts from one of her arguments).

I certainly did *not* enter the judging room in session 9 offering to help with difficult judging. I thought the judging should be completed. The experimenter had shown me a one-line transcript when he had entered my office to get the judging set, and I thought that after some minutes the judging must have been done. In fact the experimenter had got more material from the subject after the session. Further, I do not consider that I exerted any influence to bias the subject towards choosing the correct picture, but this is only a matter of opinion; I think I didn't, Blackmore thinks I did (but then she didn't express this view at the time, whereas she most certainly did in session 7).

We have already seen that Blackmore had this session rejudged, the result didn't come out right for her, and so she suppressed the evidence (in the 1987 report). Other elements of this session need stressing. First, it was certainly wrong on my part to enter the judging room. I should not have done so and this was unequivocally an error. Technically there is a faint possibility of sensory cueing given that a randomizer might have picked up some cue from the target envelope (if these target envelopes are all the same, which they should be, then this isn't of course possible. This is a key point to be returned to). Having said that, an extraordinary mistake was made by the experimenter. He added the scores up wrongly so that the target was erroneously placed first; on the item-scoring it should have been placed second. So much for the effectiveness of my biasing. We only discovered this later (I was not sitting with the experimenter and didn't see him scoring the data at the time) and obviously rechecked all his other additions (Blackmore was present at this, with Trevor Harley, although she doesn't mention this to the reader, giving instead the ambiguous account that I did the recheck. True, but her witnessing it is obviously of note). Note that this means that the independent judge Blackmore used rated the picture *higher than it was in the original session*, since the picture was actually scored second in the original session and only ranked and rated first through an error by the student experimenter.

Blackmore also misinforms the reader of the destruction of the B-envelope in two ways. First, Trevor Harley did not remember my telling him about doing this; he was actually present when I did it, before session 9, and I told him what I was doing and why. He did not remember this when Blackmore made her observations known to him, and indeed *I* didn't remember this at first; I told him that I had destroyed it and was beginning to explain why when he interrupted and told me that he had been there when I had been doing this, throwing away the torn-up envelope in a litter bin. Why these errors of immediate recall exist is uninteresting; probably under stress, Trevor Harley simply forgot this at first.

Further, in her 1987 paper Blackmore has developed an amnesia for why this was necessary. It is instructive to compare the 1987 and 1979 papers at this point.

1987 version: 'it is hard to see why a bent card (*sic*) would need to be removed from the pile since it would still be of unknown contents until it were used and destroyed'.

1979 version: 'It is of course important that these envelopes are not capable of being distinguished, as a bent one might be' (p. 13).

(Note: please refer to the Note at the end of this paper.)

Obviously the comprehension in 1979 has had to be swept aside to support the more sceptical tone of the 1987 version. The reason why removing any identifiable (e.g., bent) envelope from the randomization pile is simple and obvious: the contents *might* be identifiable *if* the randomizer had bent the envelope in taking it from the replacement pile of B-envelopes. One would then know the bent card was a B. With the Rand table random selection from the randomization pack of envelopes this should not be a problem, but if the target envelope so chosen was the one before or after the bent envelope then the bent one might be wrongly picked out (it would come more easily to hand than the one below it, obviously) by a subconsciously motivated error if the randomizer had selected the picture set to be used and knew that the B picture was a pleasant one. This possibility probably only exists in theory, but it's still one which should be protected against. Blackmore understood this perfectly well in her 1979 report. But history must be rewritten to support her case; we already have plenty of evidence of that.

Overall, we have major evidence of Blackmore's misreporting of this session and the events surrounding it. She suppresses evidence of judging she undertook; she now appears not to understand an obvious experimental precaution, the shedding of a flawed randomization envelope; she misreports Trevor Harley's presence at this time; she does not detail her presence at an empirical recheck of experimenter error; and so on. The reader does not need me to recap everything.

There is a final point of importance about this session. It was one in which I did not expect, nor wish, to see a direct hit. This session was of 15 minutes duration which, by the theory of altered-state psi-optimization which I was working with and was committed to, should not be long enough to enhance psi. Indeed, *this condition was actually built in as a baseline control* for comparison with the 30-minute duration sessions; the 15-minute sessions were predicted to give chance results, the 30-minute sessions were predicted to give significant positive scoring. In fact, of the first four experienced subjects who had done 15-minute sessions, three had scored direct hits. These were most unwelcome; we had predicted chance results here! I am emphatically *not* the kind of rank empiricist who thinks that significant psi results are wonderful no matter what form they take; anyone who has read anything with any theoretical or methodological elements which I've written knows this. What kind of fraud fakes significant results in an experimental condition in which he has specifically predicted chance results?

The 'Fraud Hypotheses'

Note to begin with there's no such thing even judging by Blackmore's papers; she actually presents a whole gamut of them, some of which even contradict each other. This 'hypothesis' can explain *anything*. But as it stands, Blackmore is impaled on a very unfortunate dilemma. Her observations support (so she claims) a possible fraud operating at the level of the selection of targets by the randomizer. Randomizer effects on the judging are ruled out as any kind of general explanation by the fact that my mistake in session 9 was a one-off occurrence; it never occurred before and it never occurred again (for any sceptic doubting this, you have a nasty problem: delete this session from Blackmore's observed data and the effects she reports are still significant). So she tries to spin a story about selection of targets which involves dovetailing with subject response bias, and proposes empirical tests to check for this. There is a fatal problem with this: her own observations show that such a fraudulent manouevure *could not possibly explain the result observed.* Why? Just look at the data for the sessions in which I acted as *experimenter* (and never left the experimental/judging room). These have the lowest mean rank-sum of any data subset for the whole batch of data. Yet, Blackmore checked for the possibility of experimenter fraud at the level of randomization. This would have to be achieved by a force-selection ploy; manipulating the randomization deck. The kind of trivial imbalance she actually found would be worthless to an experimenter trying this; the maximum effect (forcing subjects always to choose A/B pictures when there were excess A/B units in the deck) would be well below the actual scoring level (not that one could actually get subjects to do this anyway. One can check the *empirical* maximum effect by a simple S–R analysis which shows an effect size of zero). No, it would be necessary for the experimenter to replace the randomization deck—e.g., with a deck of entirely A's, or B's, or whatever. This way the experimenter would know the target identity and could bias judging enough to get a significantly higher scoring rate than the 25 per cent chance rate. Blackmore checked for this, looking at envelope markings, and she found that the randomization deck was *not* subsituted for in this way. This is clear evidence *against* experimenter fraud (as opposed to agent fraud), although for obvious reasons Blackmore doesn't want to spell this fact out. Because, having proposed the agent-randomizer fraud scenario, she conflates this with an experimenter effect relating to *my presence*, in any role in the experiment. But the occurrence of high scoring when I act as experimenter tells strongly *against* her agent-randomizer fraud story, since with the latter one would only expect significant scoring when I was agent-randomizer. Blackmore doesn't want to spell *that* out either.

Let's focus this point: Blackmore shows that scoring was significantly higher in my presence than in my absence. She claims to have evidence to show that there could have been manipulation of target selection at the level of agent-randomizer. But she also has evidence against fraud at the level of experimenter, which she obviously doesn't want to present as such, although she is not in this instance prepared to suppress it. Yet the actual scoring observed shows that it does not matter what role I play in the experimental sessions! Her agent-randomizer story is not going to explain this, and indeed the overall scoring suggests that no such story can be supported when scoring is equal in both agent-randomizer and experimenter roles. Blackmore simply gets confused at this point, as her further arguments show.

The first illustration of how confused she gets, and how desperate she is for supporting evidence, comes in her section 'Further Hypotheses'. She claims here to show that 'there is evidence that scores were higher when Sargent took part both in the few sessions observed during this visit *and in published data*', and cites the Sargent *et al.* (1980) paper in support of this claim. This is utter nonsense. To be fair to Blackmore, she has lifted this claim from an utterly unreliable source (Parker & Wiklund), but then she should have checked it. Nowhere in that paper is any data presented which would have permitted an analysis of sessions in which I was present *versus* those in which I was absent. Blackmore is citing a quite literally fabricated analysis. I know this with complete certainty *because I was present at every one of the 32 sessions in that experiment.*

If this isn't culpable enough (matters of this gravity deserve a damn sight less cavalier treatment than this), Blackmore *could* have quoted another finding from the same paper which would have made life very difficult for her. In that experiment I scored significantly above chance as a *subject* in my eight sessions. In this role it is obviously impossible to do anything to detect the target other than fixing the randomization deck (and don't forget she checked for this very possibility). What's more, she could also have cited the Sargent *et al.* (1982) article which *does* give scoring data for sessions run in my presence and absence; but then the difference was utterly insignificant ($P > 0{\cdot}20$) so this wouldn't suit her story at all. The same analysis was also given for the Sargent *et al.* (1981) paper and again the results were not significant, but there was one experimenter effect found there: Trevor Harley obtained significantly higher scoring as an experimenter than as a sender. Another null result which doesn't suit Blackmore's books. And in the two papers (1981, 1982) another result is found concerning an experimenter effect on my part which is highly important and not discussed at all by Blackmore, even though at last she could have found a significant experimenter effect to talk about. In both these experiments, a *predicted* improvement in scoring within the session was found to be highly significant *only in my absence from any role in the sessions*, and at chance when I participated, with the difference being significant in both cases. This isn't to be dismissed as some recherché post-hoc effect; it is theoretically crucial to the altered-state noise-reduction model of psi-optimization and it had been observed before which is why it was a predicted effect in both experiments (see Sargent, 1980, Experiment V). Blackmore *could* have found a lot of evidence to show that her story wouldn't hold water; instead she has resorted to a second-hand, unchecked, fabricated pseudo-analysis.

Another point which is crucial to the entire argument is the nature of Blackmore's fraud hypotheses. She claims, and she has *no independent evidence whatsoever* to support this, that her hypotheses of fixing the randomization deck *et al.* were arrived at before she made the observations which allegedly support them. I question this. The most obvious reason is their sheer stupidity. Consider instead a form of agent fraud which is much simpler and far less detectable than anything Blackmore suggests; all the agent has to do is to take an envelope of his choice from a replacement deck (thus knowing what it is) and simply takes one off the randomization deck and puts this on top of the replacement deck. In this case the numbes of envelopes in all decks would stay the same during the sessions. Afterwards all one has to do is to take the top envelope back off the replacement deck and put it in the randomization deck. Simple, easy to remember, undetectable, and none of this nonsense about changing numbers of envelopes in decks, spare envelopes, and all the rest of it. My counter-hypothesis is simply that Blackmore observed certain trivial and random errors and *then* made up hypotheses to go with them. I have parsimony on my side at least. Blackmore might reply that such a simpler method wouldn't be needed if checks weren't being made, but then if she is being truthful for the reader in expressing the scepticism she says she showed in the laboratory it is obvious that any intelligent fraud would be very careful indeed about doing anything fishy when she was around. The smartest thing to do, of course, would be to let everything go on normally and, if the results were poor, blame the experimenter effect. I don't care whether people think I'm a fraud or not, but I object very strongly to anyone believing that I might be a stupid one (not to mention being Machiavellian enough to fabricate good scores in a condition where I had predicted chance results and where significant results would have been a major embarrassment).

But, finally, Blackmore actually gives the game away herself. She actually admits that the matter of using extra envelopes was something which she 'thought of some days later'—i.e., *after* she actually found them. This is an open admission that she merely capitalized on random and trivial errors and built a fairytale around them.

The Errors

Blackmore makes much about wanting a listing of reasons for the errors made (spare envelopes, errors in the randomization deck, and the like). This is pure polemic. After all, both Trevor Harley and I have letters from her which were sent after her visit which state explicitly that she accepted our explanations for the errors found and wished to apologize for certain aspects of her behaviour. Now it seems she has developed an amnesia for all this just as she has developed an amnesia for why identifiable envelopes need removing from a randomization deck. Why did the student experimenter make a simple addition error involving half-a-dozen numbers in session 9? I don't know, but I do know that sceptics would have gone to town if *I* had made this kind of mistake even though in his case its innocence is presumably accepted without demur. Likewise I don't know how the errors in the randomization deck occurred (although these are wholly trivial), I don't know why Trevor Harley initially forgot witnessing my destruction of the bent envelope, and I don't know why Blackmore so bungled the independent judging of session 9. But I do know that Blackmore's claims that the roots of these errors could be ascertained by re-analysis is just nonsense and the reader should not be misled by this claim. It is obvious that none of the trivial errors could be so understood. Others have been accounted for. For example, the spare envelopes came from an experiment with a local medium (which Trevor Harley wasn't involved with, so there was no reason why he should have known about them) which I had been involved with. The presence of a couple of erroneous envelopes in one of the replacement piles was probably due to an error on my part in making up the original stock (doing 100+ at a time such errors can occur)—irritating but wholly trivial as an S–R analysis showed.

Re-analysis

Blackmore would have the reader believe that '(s)everal informal requests' were made for data for re-analysis. There were in fact two. One was from her, and obviously this is a non-starter. Having re-analyzed our data when she was at Cambridge, what can she possibly want now? Answer: there has been time to work with anything she recorded about the original data and to spin some new 'hypotheses' from the data which can then be allegedly supported by a re-analysis of it. Clearly, one would have to be made to allow this. The other was from Parker, a researcher who has published and circulated a string of libellous attacks on my work and integrity. Martin Johnson's PA Committee submitted the Parker-Wiklund paper (the least venomous of the three) to two jurists who both confirmed its libellous nature. Add to this the fabricated analysis Parker & Wiklund describe in their paper, *and* a wholly erroneous re-analysis Parker conducted and cited in his 1981 paper (it's in the original draft; I caught it in the proof stage and weeded it out) and the reader can see that I have been asked for data by an accomplished libeller who is incompetent into the bargain. Not only will I not permit this, but I refuse to permit the possibility of such an individual getting his hands on any of my data to publish still further calumnies of this type. And in any event Blackmore can have extracted all kinds of trends from the original data and now flourish these as hypotheses to be tested!

We can add to all this (as if anything else needed to be added) the fact that post-hoc analysis is an utter waste of time for an innocent researcher anyway. The Pratt-Woodruff experiment has been through 25+ years of this kind of nonsense without being defeated but you won't find any sceptic admitting that the sceptical case has failed here. Blackmore claims that if certain effects were not found as predicted then the fraud hypothesis would lose most of its force; history shows that this is just self-serving untruth. I'm not spending decades defending my results against ever-changing, ever-reformulated reanalyses. *You can never beat the fraud hypothesis*; it is *totally* unfalsifiable and letting oneself in for this kind of business is a form of masochism I'm not partial to. I stand by all the data reported; it is a matter of utter indifference to me whether Blackmore or anyone else in parapsychology believes me or not. I know the results were real because I was there and *my experience* tells me so. If I learned one thing in parapsychology, it is that results and statistics and data never changed anyone's mind about anything; experience is the only arbiter.

Whose Integrity is in Question?

The reader should be informed of the kind of behaviour which Blackmore has indulged in these past years. After the multiple deceptions of her visit, she wrote to both Trevor Harley and myself stating that she accepted our explanations for the errors which had occurred. She also accepted the fact that she was not going to be welcomed back with rather better grace than her 1987 paper suggests. Having done this, she then proceeded behind our backs to spread defamatory rumours and insinuations of fraud. My colleague Trevor Harley has collected testimony to this from academics, and when Martin Johnson's PA Committee investigated the rumour network, even they concluded that Blackmore was 'probably' responsible for instigating the rumours. We know that she *certainly* did so, even though she denies it.

Add to this Blackmore's suppression of multiple points of relevant evidence and her citation of a fabricated pseudo-analysis (not fabricated by her, I must point out) and I consider that we have a picture of behaviour which is below that acceptable in academic parapsychology. Johnson's PA Committee also viewed her rumour-spreading as unethical (Blackmore hastens to inform the reader that the PA didn't like my not replying to a letter from their President, delivered months late, but she wants to keep very quiet about *this* censure). The reader may judge for himself. For my part, if we add Parker's multiple libels to all this, I know it's less than other researchers like Targ & Puthoff have had to suffer, but it's enough to anger me. I do not intend to, and I shall not, address this issue again.

Note: Due to a misunderstanding on the part of the *Journal* editor. I have not been furnished with a copy of the article by Blackmore as it appears in this issue: I am working from a previous draft. I understand that the passage referring to this matter has been revised in the final draft of Blackmore's paper and my comments must be considered accordingly, keeping in mind that the version I have of Blackmore's paper is one which has been circulated by her prior to publication.

REFERENCES

Blackmore, S. (1986) *Adventures of a Parapsychologist.* Buffalo, NY: Prometheus Books.

Parker, A. (1981) In defense of introverts. *EJP* **3**, 373–380.

Sargent, C. L. (1980) *Exploring Psi in the Ganzfeld.* New York: Parapsychology Foundation.

Sargent, C. L., Bartlett, H. J., Moss, S. P. (1982) 'Response structure and Temporal incline in Ganzfeld free-response GESP testing'. **46**, 85–110.

Sargent, C. L., Harley, T. A., Lane, J. and Radcliffe, K. (1981) 'Ganzfeld psi-optimization in relation to session duration'. *RIP 1980.* 82–84.