Journal of the Society for Psychical Research Vol. 65, pp.122-137 (2001)

EXPERIMENTS ON THE SENSE OF BEING STARED AT: THE ELIMINATION OF POSSIBLE ARTEFACTS

by RUPERT SHELDRAKE

INTRODUCTION

The feeling of being stared at from behind is very well known. Surveys in Europe and North America have shown that between 70 and 97% of the population have experienced it (Braud, Shafer & Andrews, 1990; Sheldrake, 1994; Cottrell, Winer & Smith, 1996).

For many years this phenomenon was surprisingly neglected by psychical researchers, and experimental investigations were few and far between. Nevertheless, most studies gave statistically significant positive results indicating that people really could tell when they were being looked at from behind (for reviews see Braud, Shafer & Andrews, 1993a; Sheldrake, 1994). Recent studies have also given significant positive results.

Two kinds of experiment have been carried out. In the first, in a randomized series of trials the subjects were looked at directly, or not looked at, and for each trial guessed whether they were being looked at or not (Sheldrake,1994, 1998, 1999, 2000a). Their guesses were either right or wrong. (In this context the term "guess" is used for want of a better way of describing the process the subjects employed in trying to detect the lookers' influence.)

In the second kind of experiment, subjects were viewed through closed circuit television (CCTV) by a looker in a different room, and could not have received any clues about when they were being looked at through normal sensory channels.. In these CCTV experiments, the subjects did not have to make any conscious guesses; their reactions were measured automatically by recording their galvanic skin response (GSR), as in lie-detector tests (Braud, Shafer & Andrews 1990, 1993a, 1993b; Schlitz & LaBerge, 1994, 1997).

A great advantage of simple experiments in which subjects make conscious guesses is that they enable many more people to take part in this research than the CCTV method. They are also closer to the real-life phenomenon, and permit a range of investigations of different variables that affect the sensitivity of subjects or the effectiveness of starers.

These experiments have again and again shown a characteristic pattern whereby the scores in the looking trials are very significantly above the chance level, whereas in the control not-looking trials the scores are not significantly different from chance (Sheldrake, 1998, 1999, 2000a).

This pattern makes sense if people really do have a sense of being stared at from behind. The sense would be expected to operate when they are indeed being stared at. By contrast, in the control trials, the subjects are being asked to detect the absence of a stare, which is a unnatural request with no parallel in real-life conditions, so it is not surprising that their guesses were at chance levels. This characteristic pattern also implies that the results of the trials are not simply a matter of cheating, subtle cues, implicit learning or errors in recording the data. These possible sources of error should have affected scores in *both* looking *and* not-looking trials.

Many people familiar with the field of psychical research, proponents and sceptics alike, find it difficult to believe that a seemingly 'paranormal' phenomenon can be investigated meaningfully by such simple and inexpensive methods, and that these experiments can give repeatable positive results. It all seems too good to be true, and arouses the suspicion that more or less subtle artefacts must underlie this effect. If so, what could they be?

The most important potential problems are as follows:

1. Peeping or peripheral vision. This seems unlikely because it is not possible to see someone sitting directly behind, and if the head were turned sufficiently to do so the movement of the head would be obvious. Nevertheless, this possibility needed to be tested experimentally, and in one of the experiments reported in this paper I did this by comparing the performance of subjects with and without blindfolds. If peeping or peripheral vision were involved, the use of blindfolds should reduce or eliminate it.
2. Auditory or olfactory clues. The subjects might hear the lookers moving their heads, or hear differences in their breathing, or hear paper rustling when they turned away, or even detect different smells depending on which way they were facing and breathing. These possibilities have been tested in a series of experiments in which lookers and subjects were separated by closed windows. The fact that there was still a still a significant positive effect showed that sounds and smells could not explain the phenomenon (Sheldrake, 2000a). The CCTV experiments also eliminated the possibility of such sensory clues.
3. Implicit learning. In trials in which feedback is given, subjects could learn to respond to subtle sensory clues or to any patterns present in the trial randomization. If so, then these forms of learning should not take place when subjects are not given feedback. One of the experiments reported in this paper was designed to investigate the effects of feedback by comparing the performance of subjects with and without it.
4. Cheating. The lookers might whisper or in some other way signal to the subjects whether they were looking or not looking. This possibility was tested and eliminated by separating lookers and subjects by closed windows (Sheldrake, 2000a), and it was also eliminated in the CCTV experiments. .
5. Faults in recording the responses. The lookers could have made mistakes in writing down the subjects' responses. This is a general problem with any kind of research in which recording is done by human observers, rather than a specific

problem with staring experiments. Although careful supervision of the lookers can reduce this possibility, perhaps it cannot eliminate it altogether. On the other hand, in experiments such as these, recording errors would not be expected to have a differential effect in test and control trials, such as those repeatedly observed in experiments of this kind. In the CCTV trials, the records of GSR were made automatically and hence would not have been subject to human recording errors.

6. Experimenter effects. These could play a part when experimenters themselves serve as lookers, especially if they expect their stares to be ineffective. Their negative expectations might influence the way they stare, for example by reducing their concentration. Under these conditions experimenter expectancy effects are not only probable, but have actually been shown to occur, as discussed below.

Experimenter Effects

Although most experiments both with direct staring and through CCTV have given significant positive results, a few have failed to detect the staring effect, notably those carried out by sceptics when the sceptics themselves acted as the lookers. Even sceptics have obtained significant positive results when students served as lookers.

In initial experiments on the sense of being stared at carried out in the laboratory of Richard Wiseman, students acted as lookers. But in these experiments, instead of looking at a single subject, the lookers were looking at two subjects, thus dividing their attention (Wiseman & Smith, 1994). Nevertheless, in a series of CCTV experiments there was a significant difference in the subjects' skin resistance in the looking and not-looking trials ($p<0.04$).

Wiseman & Smith (1994) tried to explain this positive result as an artefact. They found that in their randomization, more looking trials preceded not-looking trials than vice versa. They argued that this could have given rise to an artefactual positive result if subjects' galvanic skin resistance (GSR) declined throughout the session as they became more relaxed. Apparently they did not examine their data to see if this was in fact the case, but concluded anyway that their positive result was "almost certainly" artefactual (Wiseman & Smith, 1994). They recommended that in future research of this kind, rather than truly random sequences, counterbalanced sequences with equal numbers of staring and not-staring trials should be used to avoid possible artefacts of this kind.

I asked Wiseman and Smith if I could examine their data to test their hypothesis. They told me that some of the data were inaccessible, but kindly supplied me with the data for 17 out of 30 subjects, which I examined to see if in fact GSR had declined throughout the sessions, as they had speculated. In 10 cases it declined, while in 7 it increased. If a general trend of decreasing GSR gave rise to artefactual positive scores when more looking trials preceded not-looking trials, in the subjects where GSR increased there should have been an opposite effect. I found that this was not the case. At least for this available subset of the data, the facts do not support Wiseman and Smith's speculation.

The positive results in this CCTV experiment were obtained with students as lookers. In subsequent CCTV experiments in Wiseman's laboratory, the experimenters themselves acted as lookers, and then there were no significant positive effects (Wiseman et al., 1995). Under these conditions there could well have been scope for experimenter effects, biassing the results in the direction of the experimenters' expectations. Such experimenter expectancy effects are well known in psychological and parapsychological research (e.g. Rosenthal, 1976; Palmer, 1989a, b).

The possibility of experimenter effects in staring experiments has been tested directly by Wiseman & Schlitz (1997), who jointly carried out a CCTV staring experiment in which half the subjects were tested with Schlitz as looker, and half with Wiseman. As on previous occasions, Schlitz obtained significant positive results (Schlitz & LaBerge, 1994, 1997), while Wiseman's were non-significant.

The implications of these experimenter effects are not symmetrical. A sceptic could well obtain a non-significant result in accordance with his negative expectations, for example by not concentrating on the subject, but someone with positive expectations could not cause subjects to obtain positive scores by any normal means when all possibilities of sensory information transfer were excluded, as they were in these experiments.

Recently Colwell, Schröder & Sladen (2000) carried out a staring experiment following similar methods to my own, using randomized sequences I had published on my world wide web site (www.sheldrake.org). In their experiment, the lookers and subjects were separated by a one-way mirror. As in my own tests, in a randomized series of trials, the looker either looked at the back of the subjects' necks, or looked away and thought of something else. The subjects guessed after each trial whether they had been looked at or not.

In the first experiment, Schröder, a graduate student, was the looker. Over nine sessions in which the subjects were given feedback as to whether their guesses were correct or not, the results were positive and statistically significant ($p<0.001$). The pattern of results was very similar to that in my own experiments (Sheldrake, 1998, 1999, 2000a), with a highly significant excess of correct guesses in the looking trials, and guesses at chance levels in the control trials, when the subjects were not looked at. There was, moreover, an increasing accuracy as the subjects were tested repeatedly, with a significant linear trend ($p < 0.003$).

Like Wiseman & Smith (1994), Colwell et al. (2000) attempted to explain their positive result as an artefact of the randomization procedure. They argued that rather than supporting the possibility that people really can feel stares, their participants' positive scores were an artefact that arose from "the detection and response to structure" present in my randomized sequences. And indeed the sequences they took from my world wide web site were not "structureless". Ironically, this was the case because I adopted the recommendation of Wiseman & Smith (1994) to use counterbalanced sequences with equal numbers of looking and not-looking trials.

The crux of Colwell et al.'s argument was that because of the deviations from "structureless" randomness in my sequences, participants could have implicitly learned to detect patterns, for example that there was a relatively high probability of an alternation after "two of a kind". But they offered no evidence that their participants in fact learned to follow such rules. They could have examined the trial-by-trial scores to see if there actually was an excess of successful guesses after two of a kind, or after any other pattern they chose to postulate. Instead, they offered no more than a speculation that this might have been the case.

The arguments of Colwell et al. (2000) were reiterated in a simplified form in an article in *The Skeptical Inquirer* by Marks Marks & Colwell (2000), who failed to mention a fundamental flaw in this hypothesis of pattern-detection through implicit learning. The problem is this. Implicit learning should in principle enable participants to improve equally in looking *and* not-looking trials. But this is not what happened. Significant improvements occurred *only* in the looking trials (Colwell et al., 2000). So how could implicit learning work in looking trials, but not at all in not-looking trials?

Colwell et al. (2000) recognized this problem, but could only suggest that participants may have "focused more on the detection of staring than non-staring episodes." This begs the question. The subjects *must* have selectively detected when staring trials were happening, otherwise their scores would not have been above chance levels and shown such an improvement in successive sessions. But this effect could well have occurred because they really could detect when they were being stared at.

Colwell et al. (2000) did a second experiment to test their hypothesis using a "structureless" randomization procedure and this time obtained non-significant results. But in their discussion, as in Marks & Colwell's (2000), they omitted to mention that in their second experiment there was not only a different randomization procedure, but also a different looker, David Sladen. As one of the proponents of the pattern-detection hypothesis, Sladen was presumably expecting a non-significant result. His negative expectations may well have influenced the way in which he stared at the participants. It would be interesting to know if Sadi Schröder, the graduate student who acted as starer in their first experiment, was more open-minded about the possibility that people really could detect stares.

Thus the difference between the results of Colwell et al.'s first and second experiments could well be have been due to an experimenter effect, rather than to differently randomized test sequences.

Colwell et al. (2000) and Marks & Colwell (2000) used the results of this second experiment to suggest that my own findings in staring experiments were an artefact due to implicit learning of structures present in the counterbalanced randomized sequences. If my experiments had involved feedback, as required by their hypothesis, this criticism might have been relevant. But this is not how the tests were done.

In more than five thousand of my own trials, the randomization was indeed "structureless", and was carried out by each starer before each trial by tossing a coin (Sheldrake 1999, Tables 1 and 2). The same was true of more than 3,000 trials in German and American schools (Sheldrake 1998). Thus the highly significant positive results in these experiments cannot be "an artifact of pseudo randomization", as Marks and Colwell (2000) suggested.

When I developed the counterbalanced sequences, I changed the experimental design so that feedback was no longer given to the subjects. Since the pattern-detection hypothesis depends on feedback, it cannot account for the fact that in more than 10,000 trials without feedback there were still highly significant positive results (Sheldrake 1999, Tables 3 and 4).

Finally, in another recent paper in *The Skeptical Inquirer,* Baker (2000) reported that in a staring experiment (through a one-way mirror) with himself as the starer the results were non- significant. Baker made no secret of his sceptical attitude and indeed regarded his experiment not so much as a test as a "demonstration" of the non-existence of an ability to detect stares. In addition to the likelihood of a strong experimenter effect, his experimental design was seriously flawed. His instructions to his subjects were confusing, ambiguous and (at least in their published form) contained serious errors, as he now recognizes (Sheldrake, 2000b).

By contrast with these experiments carried out by sceptics, the positive and highly significant results in the far larger number of experiments carried out by other investigators and by myself have involved hundreds of different people acting as lookers, with no selection for sceptical or non-sceptical attitudes. Overall, there was an extremely significant positive effect ($p< 1x10-15$) (Sheldrake, 1999), indicating that people really can tell when they are being looked at from behind.

The Effects of Blindfolding Subjects and Giving Them Feedback

In this paper I describe the results of experiments I carried out in a school in North London to examine the effects of blindfolding the subjects, compared with not blindfolding them, and giving them feedback, compared with not giving them feedback.

I also report a series of experiments in a school in Ireland in which all subjects were blindfolded and given no feedback. The Irish experiments also examined the effects of different kinds of relationship between lookers and subjects, in particular comparing unrelated pairs of children with pairs of siblings and with twins. The twin studies are of particular interest in view of the wealth of anecdotal evidence (e.g. Eason, 1994) and scientific studies (e.g. Dossey,1997) that suggest that monozygotic twins are closely linked, and may be subject to unexplained influences from each other.

The results confirm that simple experiments on the feeling of being stared at can give remarkably consistent results, even when carried out as projects by schoolchildren, and illustrate that such experiments can be used to explore new questions.

METHODS

My experiments were carried out in February and March, 1997 at University College School Junior Branch (UCS), a boys' school in Hampstead, London. Each experiment took place with a different class, either in the fourth form (age 10-11), or in the third form (age 9-10). The experiments were carried out in the school science laboratory and were supervised by myself and the class's science teacher, either Mr Mark Albini or Mr John Hubbard. Before the experiment began, I gave a brief introductory talk explaining and demonstrating the procedure.

The boys worked in pairs, one (the subject) sitting with his back towards the other (the looker). The distance between them was

1 to 2 metres. They sat in places where there were no reflective surfaces (such as mirrors or windows) that could have enabled the subject to see the looker's reflection. Each pair of boys sat in a different part of the room, and proceeded with the trials at their own pace.

In a set of 20 trials, in a random sequence, the looker either looked at the back of the subject or looked away, and was instructed to think of something else. The random sequence was set out on previously prepared sheets, with 24 different random sequences of looking and not-looking trials, compiled on the basis of standard random number tables. These sheets were given to the lookers only after the subject was in place and unable to see the sheet. Following the suggestion of Wiseman & Smith (1994) that trial sequences in staring experiments be counterbalanced, 12 of these 24 trial sequences were the inverse of the other 12.

The looker indicated to the subject when a trial was beginning by a click, made with a mechanical clicker, and the subject then guessed whether he was being looked at or not, saying out loud "looking" or "not looking". The looker recorded the result on the instruction sheet. In trials in which feedback was given, the looker then told the subject whether the answer was correct or not.

Usually subjects indicated their guess within 10 seconds, but if they had not done so already were asked to do so after 20 seconds. The procedure was therefore quite fast, and most pairs could complete 20 trials in 10 minutes or less.

The blindfolds worn by subjects were kindly supplied by Virgin Atlantic Airways, and were of the type widely used by air passengers in order to sleep on planes. They are held on by elastic bands that go round the back of the head and block out all sideways peripheral vision, although light can sometimes leak in under the blindfolds right next to the nose. These blindfolds effectively eliminate any possibility of the subjects seeing what is going on behind them.

Experiment 1: The effects of blindfolds

The experiments to test the effects of blindfolds were carried out by classes 4B and 4W. Each class was divided into two groups, 1 and 2, consisting of 4 or 5 pairs of boys. The experiments took place in two phases, and within each phase there were two sessions. In the first phase of the experiment, the subjects in group 1 wore blindfolds, while those in group 2 did not. In the first session, each subject completed a set of 20 trials, and then for the second session the looker and subject changed places and carried out a second set of 20 trials. In the second phase of the experiment, the subjects in group 1 did not wear blindfolds, while the subjects in group 2 did so. Again, there were two sessions in which first one and then the other member of each pair took his turn as subject. In the case of class 4B there was sufficient time for a third phase, in which the pairs in group 1 again used blindfolds, and those in group 2 did not. In all cases, subjects were told after each guess if they were right or wrong, in other words they received feedback.

Experiment 2: The effects of feedback

In the experiments on the effects of feedback, all subjects wore blindfolds in all trials. These experiments were conducted with classes 4S and 3E.

As in the experiments on the effects of blindfolds, each class was divided into two groups, 1 and 2. In the first phase of the experiment, the subjects in group 1 were given feedback, while those in group 2 were not; in the second phase the situation was reversed, with the subjects in group 2 receiving feedback and those in group 1 not receiving it. As before, there were two sessions in each phase, in which the lookers and subject reversed roles.

Experiments in Irish schools

In 1998, two Irish schoolgirls, Susan and Jennifer Brodigan, who are non-identical twins, carried out a large-scale project on the feeling of being stared at, following instructions I supplied. In their experiments all subjects wore blindfolds, and none received feedback.

The participants in their experiments were 11 to 18-year -old girls at the twins' own school, Our Lady's College, Greenhills, Drogheda, and also primary school children, both boys and girls, aged 9-11 from a number of different schools: Tullyallen National School; St Patrick's National School, Harestown; Presentation School, Ballymakenny; Cartown National School, Termonfeckin; and Scoil Aonghusa. In their experiments some pairs of children were related, either siblings or twins, while other pairs were unrelated.

The methods were as described above, except for the fact that instead of each looker signalling individually when each trial began, a signal was given to the whole class by means of a buzzer, so the trials in a given class were synchronized. Each looker had a different score sheet with a different random sequence of looking and not-looking trials.

The Brodigan twins kindly sent me all the score sheets from their experiments, and I tabulated the data as described below.

Analysis of the Data

The numbers of right and wrong guesses from each set of 20 trials carried out by each looker-subject pair were tabulated in three columns; "Looking", "Not looking" and "Total", enabling the total number of right and wrong guesses in each column to be obtained.

For each set of 20 trials, in each column, the data were also scored as follows:

+ if the subject made more right than wrong guesses
- if the subject made more wrong than right guesses
= if the number of right and wrong guesses was the same.

Statistical analysis was carried out in three ways. First, the chi-squared test was used to compare the total numbers of right and wrong guesses in each column. The null hypothesis was that the numbers of right and wrong guesses would be the same.

Second, the chi-squared test was used to compare the total numbers of + and - scores. The = scores were disregarded. The null hypothesis was that by chance the number of + and - scores would be equal. This method of analysis was suggested to me by Prof. Nicholas Humphrey as a way of testing if effects were broadly spread among participants, since this method gives an equal weight to each.

Table 1

Staring experiments with and without blindfolded subjects (classes 4W and 4B at UCS).
All subjects were given feedback.

A: *Numbers of right and wrong guesses (percentage of right guesses shown in parentheses).*

	Looking		Not Looking		Total	
	right	wrong	right	wrong	right	wrong
Blindfold	260 (65%)	140	202 (50.5%)	198	462 (57.8%)	338
No Blindfold	240 (63.3%)	139	208 (54.6%)	173	448 (58.9%)	312

p values

Comparison of total numbers of right and wrong guesses by chi-squared test:

	Looking	Not Looking	Total
Blindfold	2×10^{-9}	n.s.	2×10^{-9}
No Blindfold	3×10^{-6}	n.s.	1×10^{-5}

Comparison of right and wrong guesses by paired-sample t-test:

Blindfold	0.005	n.s.	0.01
No Blindfold	0.008	n.s.	0.001

B Accuracy of guesses (40 subjects blindfold (b); 38 subjects not blindfolded (nb))

	Looking		Not Looking		Total	
	b	**nb**	**b**	**nb**	**b**	**nb**
More right than wrong	28	25	19	18	24	26
More wrong than right	7	11	17	12	11	8
Equal right and wrong	5	2	4	8	5	4

p *values.*

Comparison of number of subjects more right than wrong with those more wrong than right by chi-squared test:

	Looking	Not Looking	Total
Blindfold	0.003	n.s.	0.03

No Blindfold	0.02	n.s.	0.003

Third, the numbers of right and wrong guesses were compared using the paired-sample t-test, with the numbers of right and wrong guesses for each group in each session as the paired sample. The null hypothesis was that the numbers of right and wrong guesses would be the same.

For the comparison of two sets of scores (for example the scores with and without blindfolds) 2 x 2 contingency tables were used (Campbell, 1989), with the null hypothesis that the proportions of right and wrong guesses in both sets were equal.

RESULTS

Experiment 1, With and Without Blindfolds

In this experiment, all subjects were given feedback.

As in previous experiments on the sense of being stared at (Sheldrake, 1999), subjects scored considerably above the chance level of 50% in the looking trials: 65.0% correct with blindfolds (p= 0.005 by the paired-sample t test) and 63.3% without (p=0.0008). By contrast, in the not-looking trials, their guesses were not significantly different from the chance level of 50% (Table 1A). The total scores, combining the results from the looking and not-looking trials, were also significantly above chance levels: 57.8% correct with blindfolds (p= 0.01 by the paired-sample t test) and 58.9% without (p= 0.001) (Table 1A).

Using the alternative system of scoring the results, according to which each subject is scored "+" if they more often right than wrong and "-" if they more often wrong than right, the overall scores were 24+ 11- with blindfolds (p= 0.03 by the chi-squared test) and 26+ 8- without (p= 0.003) (Table 1B).

Thus blindfolds had little effect on the subjects' performance. There was no significant difference between the scores with and without blindfolds.

Experiment 2, With and Without Feedback

In this experiment all subjects wore blindfolds.

The general pattern of results showed, as usual, that there were significantly more correct than incorrect guesses in the looking trials: with feedback 59.8% were correct (p= 0.001 by the paired sample t test) and without feedback 60.2% (p= 0.01) (Table 2A). Likewise, the subject-by-subject scores showed a significant excess of people who were more often right than wrong both with feedback (31+ 11- ; p= 0.002 by the chi-squared test) and without feedback (28+ 15- ; p= 0.05) (Table 2B). By contrast in the not-looking trials the scores were at chance levels. (Table 2A and B).

Thus giving the subjects feedback made very little difference to their performance. Overall, there was a slightly higher percentage of correct guesses without feedback (54.9%) than with feedback (53.6%), but this difference was not significant.

In general the scores were lower than in experiment 1. The subjects given feedback in experiment 2 were tested under the same conditions as the blindfolded subjects in experiment 1, that is to say they were blindfolded and given feedback, but their scores were lower (in total, 53.6% correct guesses as opposed to 57.8%; Tables 1A and 2A).

Table 2

Staring experiments with and without feedback to subjects (classes 4S and 3E at UCS). All subjects wore blindfolds.

A: *Numbers of right and wrong guesses (percentage of right guesses shown in parentheses).*

	Looking		Not Looking		Total	
	right	wrong	right	wrong	right	wrong
Feedback	275 (59.8%)	185	218 (47.4%)	242	493 (53.6%)	427
No Feedback	283 (60.2%)	187	233 (49.6%)	237	516 (54.9%)	424

p *values.*

Comparison of total numbers of right and wrong guesses by chi-squared test:

	Looking	Not Looking	Total
Feedback	3×10^{-4}	n.s.	0.03
No Feedback	1×10^{-5}	n.s.	0.002

Comparison of right and wrong guesses by paired-sample t-test:

	Looking	Not Looking	Total
Feedback	0.001	n.s.	n.s.
No Feedback	0.01	n.s.	n.s.

B Accuracy of guesses (46 subjects given feedback (f); 47 subjects not given feedback (nf))

	Looking		Not Looking		Total	
	f	**nf**	**f**	**nf**	**f**	**nf**
More right than wrong	31	28	15	19	21	24
More wrong than right	11	15	24	21	16	17
Equal right and wrong	4	4	7	7	9	6

***p** values.*

Comparison of number of subjects more right than wrong with those more wrong than right by chi-squared test:

	Looking	Not Looking	Total
Feedback	0.002	n.s.	n.s.
No Feedback	0.05	n.s.	n.s.

Do Subjects Improve with Practice?

When subjects are given feedback, there is the possibility that can improve their sensitivity with practice. Since all subjects in the experiments of series 1 were given feedback, such an effect should be detectable by comparing the results of phase I of the experiment, in which subjects were tested for the first time, with phase II, in which the same subjects were tested again. In fact, in phase II, the scores did improve: the percentage of correct guesses rose from 55.6% to 62.1%. This increase was statistically significant (p=0.02 by the chi-squared test using two-way contingency tables).

Another way of evaluating improvement is to compare the first 10 trials with the second 10 in each set of 20 trials . In the experiments in series 1, there was a slight overall improvement, from 57.2% to 59.1%, but this difference was not statistically significant.

In series 2, with and without feedback, there was no improvement in scores from phase I to phase II: the percentage of correct guesses was 53.5% in both. When the first 10 trials in each set were compared with the second 10, there was a slight decline in the percentage of correct guesses, from 54.7% to 52.1%, but this was not statistically significant.

Experiments in Irish Schools

In the experiments carried out in Irish schools by Susan and Jennifer Brodigan, all subjects wore blindfolds and none were given feedback.

The results showed the now-familiar pattern whereby scores were well above chance in the looking trials with a total of 56.7% correct, and with subjectwise scores of 114+ and 52- (p=2x10-6 by the chi-squared test). By contrast, the scores were at chance levels in the not-looking trials (Table 3). A similar pattern was apparent in all the sets of data.

There was a tendency for the scores to be higher with the children who were related (overall 55.1% correct guesses) than with those who were unrelated (overall 52.4% correct), but this difference was not statistically significant.

Table 3

Staring experiments with blindfolds and without feedback in Our Lady's College, Drogheda, Ireland.

A: *Numbers of right and wrong guesses (percentage of right guesses shown in parentheses).*

	Looking		Not Looking		Total	
	right	wrong	right	wrong	right	wrong
Unrelated						
Total Unrelated	926 (55.9%)	730	815 (49%)	849	1741 (52.4%)	1579
Related						
Siblings	163 (54.3%)	137	151 (50.3%)	149	314 (52.3%)	286
Identical Twins	93 (58.1%)	67	81 (50.6%)	79	174 (54.4%)	146
Non-identical Twins	131 (65.5%)	69	106 (53.0%)	94	237 (59.3%)	163
Total Related	387 (58.6%)	273	338 (51.2%)	322	725 (54.9%)	595
Grand Totals	1313 (56.7%)	1003	1153 (49.6%)	1171	2466 (53.1%)	2174

p *values.*

Comparison of number of total number of right and wrong guesses by chi-squared test:

Looking	Not Looking	Total
1×10^{-9}	n.s.	2×10^{-3}

p *values.*

Comparison of number of total number of right and wrong guesses by paired-sample t-test:

Looking	Not Looking	Total
0.03	n.s.	0.05

B: *Accuracy of guesses (166 unrelated subjects; 30 siblings; 16 identical twins and 20 non-identical twins).*

Numbers of subjects with more right than wrong guesses (+), more wrong than right guesses (-) or equal numbers of right and wrong guesses (=)

Group	Looking			Not Looking			Total		
Unrelated									
Total Unrelated	78+	42-	46=	51+	68-	47=	77+	65-	24=
Related									
Siblings	11+	7-	12=	11+	11-	8=	16+	11-	3=
Identical Twins	8+	2-	6=	5+	7-	4=	7+	2-	7=
Non-identical Twins	17+	1-	2=	9+	9-	2=	12+	4-	4=
Total Related	36+	10-	20=	25+	27-	14=	35+	17-	14=
Grand Totals	114+	52-	66=	76+	95-	61=	112+	82-	38=

p *values.*

Comparison of number of + and - scores by chi-squared test:

Looking	Not Looking	Total
2×10^{-6}	n.s.	0.03

Among the related children, the untwinned siblings did no better than unrelated children (overall 52.3% correct, as opposed to 52.4%). The non-identical twins did better than the identical twins, with 59.3% and 54.4% correct guesses respectively (Table 3), but these differences were not statistically significant.

DISCUSSION

Blindfolding subjects made no significant difference to their ability to tell when they were being looked at from behind (Table 1). The type of blindfolds used in these experiments effectively eliminated peripheral vision. Even if were to be argued that these blindfolds allowed for some vision by looking downward along the nose, they could not have allowed the subject to see what was going on behind them unless they turned their head around and tilted it backwards. No subjects were observed to be doing this. The fact that the blindfolds made no significant difference to the results shows that the effect detected in these experiments did not depend on visual clues.

There was also no significant difference between the scores with and without feedback (Table 2), showing that the ability of subjects to detect when they were being looked at did not depend on receiving feedback.

A puzzling aspect of the data is the fact that the scores were generally higher in experiment 1 than in experiment 2. Even when the conditions were identical, as they were for blindfolded subjects in experiment 1 (who were given feedback) and subjects given feedback in experiment 2 (who were blindfolded), the subjects in experiment 1 had higher scores than in experiment 2. Why?

This difference could just be a matter of chance. Perhaps the children in experiment 1 just happened to be more sensitive than the children in experiment 2. But there could be a more interesting reason for this difference. Perhaps in experiment 2 the participants were generally more self-conscious about their performance. Subjects received feedback in some sets of trials, while in others they did not, which could have made them anxious about how they were performing, with the effect of reducing their sensitivity. But in the absence of further research this can be no more than a speculation.

The relatively low scores in the experiments with and without feedback do not, however, alter the main conclusion that sense of being stared at is still detectable without feedback.

After completing the experiments described in this paper, I arranged for further experiments on the sense of being stared at to be conducted without feedback in schools in Connecticut, USA, following an earlier series with feedback (Sheldrake, 1998). There were highly significant positive scores without feedback, just as there had been with feedback. Tests were conducted in 8 schools. In a total of more than 5,000 trials the overall proportion of correct guesses was 55.3%. Subjectwise the scores were 149+ 74-, in other words 149 subjects were more often right than wrong, compared with 74 who were more often wrong than right (p= 5x10-7 by the chi-squared test) (Sheldrake, 1999).

These experiments confirmed that positive scores in these experiments were not dependent on feedback. But they did not did not involve the use of blindfolds. In the experiments in Irish schools described in this paper (Table 3), the subjects not only received no feedback but were also blindfolded. The results confirm that the effect detected in these experiments depends neither on visual clues nor on feedback.

The positive results of these experiments with blindfolded subjects deprived of feedback shows that they do not arise from artefacts due to visual clues. Nor are they due to implicit learning dependent either on sensory clues, or on the detection of subtle patterns in the randomization of trials.

The experiments reported in this paper do not, however, eliminate the possibility of auditory or even olfactory clues, because the lookers and subjects were in the same room and only 1-2 metres apart. However, any hypothesis that proposes that such clues were involved would have to explain why they worked only in the looking trials but not in the not-looking trials. Most conceivable clues of this kind (including deliberate attempts to cheat, for example by the looker whispering to the subject) would be expected to elevate the scores in both looking and not-looking trials. But this is not what happened.

Nevertheless, a sceptic might propose that there were clues present only in the looking trials that the subjects were intermittently and unconsciously aware of. The only way of answering this argument definitively is through experiments in which lookers and subjects are separated by soundproof barriers. I have carried out such experiments through closed windows, and again the results showed a significant positive effect, with the usual difference between looking and not-looking trials (Sheldrake, 2000a).

Moreover, in the experiments of Colwell et al. (2000), discussed in the Introduction to this paper, the subjects were looked at through a one-way mirror, which served as a barrier to auditory and olfactory cues. With a graduate student as looker, subjects scored very significantly above chance, with the usual pattern of highly significant positive scores in looking trials and non-significant scores in not-looking trials. And in most of the experiments carried out through CCTV there were significant positive results under conditions where there was no possibility of the transfer of any auditory or other sensory clues (Braud, Shafer & Andrews 1990, 1993a, 1993b; Schlitz & LaBerge, 1994, 1997). The exceptions were experiments in which sceptics themselves were the lookers, as discussed in the introduction to this paper, and may well have involved an experimenter effect of the kind reported by Wiseman & Schlitz (1997).

The pioneering research of Susan and Jennifer Brodigan suggests that twins, both identical and non-identical, perform better than non-twinned siblings or unrelated children; but further research with larger samples would be necessary to reach a firm conclusion on this question.

The sense of being stared at does not seem to be explicable in terms of normal sensory information. It must therefore depend on causal factors at present unknown to science. Some possible explanations have been discussed by Abraham, McKenna & Sheldrake (1992) and Sheldrake (1994).

ACKNOWLEDGMENTS

I am grateful to all the people who took part in these experiments, to John Hubbard and Mark Albini for helping me to do the tests in their classes, and to Jennifer and Susan Brodigan for sending me their results and agreeing to their publication in this paper. I thank the Institute of Noetic Sciences, Sausalito, CA, the Lifebridge Foundation, New York and the Bial Foundation, Portugal, for financial support.

REFERENCES

Abraham, R., McKenna, T. & Sheldrake, R. (1992) *Trialogues at the Edge of the West.* Santa Fe: Bear & Co.
Baker, R. (2000) Can we tell when someone is staring at us from behind? *Skeptical Inquirer* (March/April), 34-40.
Braud, W, Shafer, D. & Andrews, S. (1990) Electrodermal correlates of remote attention: Autonomic reactions to an unseen gaze. *Proceedings of Presented Papers, Parapsychology Association 33rd Annual Convention,* Chevy Chase, MD, pp14-28.
Braud, W, Shafer, D. & Andrews (1993a) Reactions to an unseen gaze (remote attention): A review, with new data on autonomic staring detection. *JP* **57,** 373-90.
Braud, W, Shafer, D. & Andrews (1993b) Further studies of autonomic detection of remote staring: replications, new control procedures, and personality correlates. *JP* **57,** 391-409.
Campbell, R.C. (1989) *Statistics for Biologists.* Cambridge: Cambridge University Press.
Colwell, J, Schröder, S, & Sladen, D. (2000) The ability to detect unseen staring: A literature review and empirical tests. *British Journal of Psychology* **91,** 71-85.
Cottrell, J.E., Winer, G.A. & Smith, M.C. (1996) Beliefs of children and adults about feeling stares of unseen others. *Developmental Psychology* **32,** 50-61.
Dossey, L. (1997) Lessons from twins: of nature, nurture and consciousness. *Alternative Therapies* **3** (3), 8-15.
Eason, C. (1994) *The Psychic Power of Children.* London: Foulsham.
Marks, D. and Colwell, J. (2000) The psychic staring effect: An artifact of pseudo randomization. *Skeptical Inquirer*

(September/October), 41-49.
Palmer, J. (1989a) Confronting the Experimenter Effect. *Parapsychology Review* **20** (4), 1-4.
Palmer, J. (1989b) Confronting the Experimenter Effect, Part 2. *Parapsychology Review* **20** (5), 1-5.
Rosenthal, R. (1976) *Experimenter Effects in Behavioral Research.* New York: John Wiley.
Schlitz, M. & LaBerge, S. (1994) Autonomic detection of remote observation: Two conceptual replications. *Proceedings of Presented Papers, Parapsychology Association 37th Annual Convention,* Amsterdam, pp. 352-60.
Schlitz, M. & LaBerge, S. (1997) Covert observation increases skin conductance in subjects unaware of when they are being observed: a replication. *JP* **61,** 185-195.
Sheldrake, R. (1994) *Seven Experiments that Could Change the World.* London: Fourth Estate, Chapter 4.
Sheldrake, R. (1998). The sense of being stared at: experiments in schools. *JSPR* **62,** 311-323.
Sheldrake, R. (1999). The 'sense of being stared at' confirmed by simple experiments. *Biology Forum* **92,** 53-76.
Sheldrake, R. (2000a). The 'sense of being stared at' does not depend on known sensory clues. *Biology Forum* **93,** 209-224.
Sheldrake, R. (2000b). Research on the feeling of beng stared at. (submitted to Skeptical Inquirer).
Wiseman, R. & Schlitz, M (1997) Experimenter effects and the remote detection of staring. *JP* **61,** 197-207.
Wiseman, R. & Smith, M. (1994) A further look at the detection of unseen gaze. *Proceedings of the Parapsychological Association 37th Annual Convention.* Parapsychological Association. 465-478.
Wiseman, R., Smith, M.D., Freedman, D., Wasserman, T. & Hurst, C. (1995) Examining the remote staring effect: two further experiments. *Proceedings of Presented Papers, Parapsychology Association 38th Annual Convention,* pp. 480-490.

Rupert Sheldrake

20 Willow Road
London NW3 1TJ
England