EXTRASENSORY PERCEPTION HOME-BASED EXPERIMENT REPORT

From 2014 to 2020, 230 tests of extrasensory perception were performed by two unrelated subjects, one male and one female. It should be noted at the outset, that the experimenter was also the female subject. The majority of the tests comprised randomly selected pictures from magazines, newspapers, books or photographs, concealed in envelopes. A certain number consisted of hand-drawn simple images concealed in envelopes. A certain number consisted of objects or pictures hidden in a distant location, inaccessible to the subject. A few consisted of a written story, placed in an envelope. The hidden test samples, whether pictures, objects, stories, or drawings, will be referred to as 'targets'. The subjects each received the randomly selected envelopes, or confirmation of the placement of an object. The relevant targets for each subject had been randomly selected by the other subject, from a pool of targets, selected, randomised, and known only to them. In this way, for most experiments, neither subject knew which target had been selected for the other subject.

The subject, in solitude, in their own chosen environment, (usually their own separate home) would hold the envelope, or picture the location of the target (by virtue of a specified visual trigger linking the subject to the target,*), and follow a technique for quieting the mind and entering an altered state of expanded consciousness. The conscious intention was stated, to see or perceive clearly what was hidden. During a period of time, varying from five to twenty minutes, any information perceived, was recorded, by writing or drawing. The information could be either visual, verbal or multisensory.

The pictures were cut out from a variety of sources, magazines, newspapers, books, cards, photographs. The drawings were very basic line drawings, of varied subjects. The stories were a paragraph long, in handwriting, and very simple, varied imaginary themes.

It will be appreciated that this was not a laboratory experiment. It was informal, and allowed the subjects, who were working on developing psychic and mediumship abilities, to find their own best method of extra-sensory perception, and obtain feedback on the efficacy of their attempts.

It was also important to test the hypothesis that it was possible for the subjects to have extra-sensory perception. That means, that subjects were able to obtain some accurate details of hidden objects and pictures over and above what might just be random guesswork. This was considered to be an important aspect of mediumship development, as the ability of a medium to perceive things hidden from normal perception can be tested by using actual objects in the physical world, whose identity can be subsequently confirmed through normal perception.

From the point of view of statistical analysis of the data after all the experiments had finished, in December 2021, a control set of pictures, objects or stories, were subsequently randomly selected and paired with the experimental targets, for scoring, by an independent and blinded scorer, who had no previous knowledge or interaction with the experiments. This independent scorer was a physics and philosophy university student, who had an interest in the study of consciousness.

It was hypothesised that if the recorded perceptions for the hidden test targets were all random and no different from chance guesses, then the scores for the control targets would follow the same distribution and clustering about an average as the scores for the test targets. In other words there

**The visual trigger was a small object owned by the subject undertaking the test, and placed by the other subject in the target location. So the subject undertaking the test, would ask the question mentally 'what is the picture or object placed by my 'trigger'? In the case of the female, this was a small rubber frog, and in the case of the male, this was a small toy car.*

would be no difference between the overall distribution of the scores for the test and the control. This would suggest that whatever perceptions the subject recorded could have applied to any target, in other words, they were generic, or very general and not at all specific. (that is called the null hypothesis, ie there is no significant difference according to an agreed level of what constitutes significant difference. This is normally set at the level of P-0.05, meaning that at least no more than 5 experiments out of every 100 are due to chance.) However, if the subject test scores were significantly different from the paired control scores, then that would suggest that some anomalous form of extra-sensory perception, was involved.

The subjects' hand-written recordings were typed, drawn and itemised on the computer and printed out. Each one was dated according to the date of the recording, and given a sequential number. A list of random numbers was downloaded from a random number generator program on the internet, from 1-500, in pairs. The matching test picture, story or drawing was dated accordingly and given a separate matching code number from the paired random number list, which identified it as the test, and the randomly selected control target was dated accordingly, and given the other code number from the pair, which would identify it as the control. The experimenter noted the identity (male or female) of the subject from the sequential number of the record sheet. The experimenter kept a key code grid, which identified each of the paired random numbers with the appropriate sequential number of the record. Each printed sheet of each of the experimental recordings was itemised and a grid drawn over it, so that each item recorded could be scored independently, and the score written in the appropriate column and box of the grid, under an appropriate code number heading noted and written in by the scorer, for each of the two separate targets. These targets were each placed in separate envelopes and the envelope marked with the appropriate code and sealed. Only the experimenter could know from the key sheet which was the control target and which was the test target, and the identity of the subject. And this would not be revealed until the end of all the scoring. For each experiment, the subject's printed record sheet and the two sealed envelopes were placed together in a larger envelope, which was labelled with the date and code number of the record sheet, and sealed. This was to ensure that each experiment would be scored and the scores noted correctly, identified clearly, over one time period, and the results placed back in the envelope and kept together with the data, without getting muddled up, and quality control and consistency could be checked.

A score sheet was compiled, which allowed for a range of criteria (parameters) to be quantified. These criteria ranged from clear identification of target to more subtle and partial identifying perceptions. * (see the score sheet below). The subject's record for each experiment was scored against both the target and the control, consecutively, by the blinded scorer, in such a way that the scorer did not know which was the target and which was the control. The experimenter had the code key to the experiment, and was not involved in this independent scoring, to avoid experimenter bias. It was considered very important to ensure that the scorer was scoring independently and consistently for each and every experiment, so that there was parity between the scores for each test and control. It was also considered important that the same scorer completed the whole set of data in as short a period of time as possible, so that the acuity and consistency of scoring was maintained.

The score sheet had been compiled by the two subjects to cover the range of different ways in which ephemeral and subtle experiences of psychic perception seem to be registered and filtered by the brain. However, because of the nature of the criteria involved, the interpretation of the individual criteria was likely to be subjective, and could vary with mood, perception, experience, ontological appreciation, conceptual acuity and consistency of the scorer. This could be partly controlled for by the scorer scoring blinded, each experiment, test and control at the same time, and the whole series within a continuous time block, and in the same place, and with the scorer in the same health

condition. Blinding could be assured by the scorer not knowing how many subjects were involved in the tests, the sex or age of the subjects, which pictures were control and which were test. These conditions were met.

*the scoring sheet used by the scorer is shown below

EXTRASENSORY EXPERIMENT SCORING

SCORE	
5	An accurate description of the whole
4	An accurate description of part (more than one unit)
3	An accurate description of one unit
2	An accurate perception of main colour
2	An accurate perception of main shape
2	An accurate perception of main texture
2	An accurate perception of main feeling/gestalt (eg happy, hot, violent, sad, fast, lazy, work, noisy, musical)
1	An associated word or image or concept (eg if it is a picture of a flower, the word might be 'garden'',fragrance', 'vase', 'growing plants','nature' 'beauty' 'bouquet')
0.5	A vaguely associated word/phrase (eg if it is a picture of a flower, the word might be 'gift', 'get well soon', 'birthday gift', 'hospital visit', 'seeds','welcome')
0.5	An accurate perception of **part** colour/ shape/ texture/ feeling/ word
0	No perceptible connection with anything pertinent

RESULTS

The results of the extra-sensory experiments were subjected to statistical testing, provided by XLSTAT, using non-parametric T-tests, as normality tests showed that the data was not normally distributed. The T-tests used, all showed a highly significant difference for the female scores between the test and control pictures, at the probability level of 0.001. However the difference for the male scores between the test and control pictures was below the level for significance, at the probability level of between 0.155 and 0.408. The probability level for significance was set at P= 0.05. This means that no more than 5 experiments out of 100 must be by chance or guesswork for a result to be deemed significant. This means that judged by this scoring, for the female subject, 1 experiment in every 1,000 would be by chance. For the male subject, judged by this scoring between 15 and 40 experiments in every 100 experiments would be by chance. For those who are not familiar with statistics or the decimal system, it should be understood, from the tables below, that the smaller the P value, the higher is the level of significance. In other words, a value of P=0.001 shows a high level of significance. A value of P= 0.1 is not significant. A value of P= 0.01 is significant. A value of P= 0.06 is not significant, although may be considered borderline significance. P=0.05 is considered just at the first level of significance.

Summary (p-values) for the Female scores

Variable\Test	Sign test	Wilcoxon signed-rank test
fem test - f control	**0.001**	**0.001**

Summary (p-values):

Variable\Test	Sign test	Wilcoxon signed-rank test
male test - m control	0.155	0.408

The two tables above show the results of the statistical analyses for the female and male experiments respectively. The figures given are the P-values, showing the statistical probability of a chance result.

It can be seen from the plots below, that there is a difference in the range of scores between the male and female. It can also be seen from both sets of scores that the highest scoring level reached for both male and female was above the highest chance scoring, and was significantly higher, at the P= 0.0001 level. This, in itself is an interesting result. However the overall scoring shows a difference between male and female in the general spread of the data, which led to the male test scoring

cluster spread around the average or mean level of the control scoring, whereas the female test scoring

cluster spread above the average or mean level of the control scoring. It was this spread that determined the difference in statistical significance, according to the standard non-parametric tests used in this analysis.

Statistical Analysis for Female Scores

Summary statistics:

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
fem test	93	0	93	0.000	14.500	2.258	2.422
f control	93	0	93	0.000	8.000	1.618	1.720

Sign test / Two-tailed test:

N+	50
Expected value	36.000
Variance (N+)	18.000
p-value (Two-tailed)	0.001
alpha	0.05

The p-value is computed using an exact method.

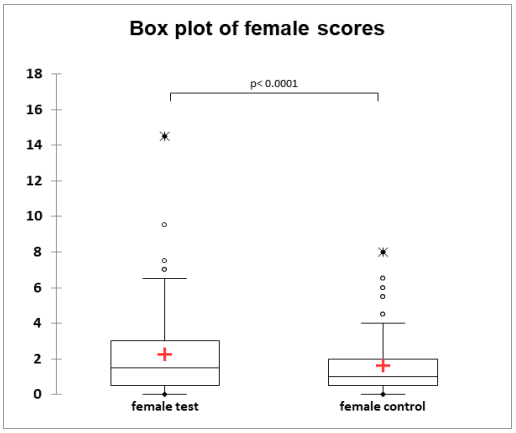
Test interpretation:

H0: The two samples follow the same distribution.

Ha: The distributions of the two samples are different.

As the computed p-value is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The plots below help to show the spread of test and control scores. There are many different ways of displaying this data. But viewing the data may help with understanding how the summary P values from the two statistical tests, relate to the spread of data and the difference in means and standard deviation. The two tests, Sign test and Wilcoxon sign-rank test are computed differently, but both tests lead to the same conclusion about significance.



Statistical analysis for the Male scores

Summary statistics:

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
male test	136	0	136	0.000	7.500	0.790	1.162
m control	136	0	136	0.000	6.500	0.849	1.104

Sign test / Two-tailed test:

N+	41
	48.50
Expected value	0
	24.25
Variance (N+)	0
p-value (Two-tailed)	0.155
alpha	0.05

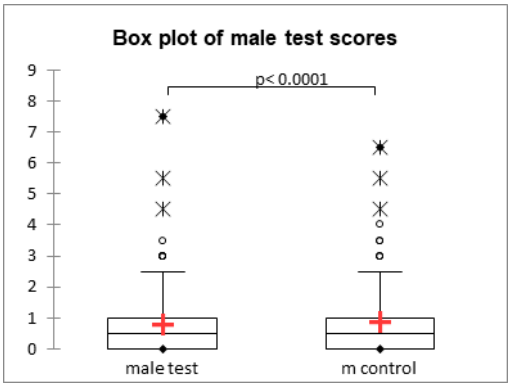
The p-value is computed using an exact method.

Test interpretation:

H0: The two samples follow the same distribution.

Ha: The distributions of the two samples are different.

As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.



Discussion

A cogent criticism for this experiment is that of bias, in that the subjects, one of whom was also the experimenter, and both of whom were actively developing mediums, would have some vested interest in supporting the hypothesis that they were capable of extra-sensory perception.

However, in defence of this criticism, both subjects were genuinely concerned to test themselves and know the truth about their abilities, even if it meant disappointment and disillusion in the conclusions they must draw from these experiments. They were engaged in this activity with exploratory integrity, rather than deliberate attempt to deceive themselves and others. So conscious bias was unlikely. However unconscious or subliminal biasing was still possible. In analysing the results, great care was taken to guard against bias. However, there was still a possibility of unconscious bias or skewing of the test results from a subliminally biased target selection process.

Ideally, in a rigorous scientific experiment, an independent source of targets and target selection and randomisation, should have been used to avoid the possibility of preferential subject choice by the two subjects, who knew each other, and therefore each other's preferences. By making the target selection themselves, they could have subliminally picked targets preferentially favouring correct guesswork by the other subject.

Proportional differences in the category selection of targets could also have played a role in the subject's test record and in the final experimental results. The female selected test targets for the male, and the male selected the test targets for the female. This was over the period of 6 years experiments. Any significant pattern in the type of targets consistently selected, may have subliminally influenced the subject's test recording. At the end of the experimental period, in 2021, two new sets of control targets were selected. The female selected the control targets for the male, as in the original test targets, but both the male and female selected the control targets for the female. This was different from the original female test targets, which had all been selected by the male. This difference was caused by a uniformity noted in the control pictures selected by the male, most of which were from recent seasonally Christmas newspapers and magazines, many featuring Christmas gifts. Because of this, it was decided by the female experimenter, who organised the data, to cut out pictures from a wider variety of sources, with no specific date or time associations. It was considered essential that no date differences could help the scorer to identify control versus test pictures. It was considered that any non-randomness in the control selection compared with the test selection, may affect the results. Therefore statistical analysis of the target selection was carried out as shown below.

The categories shown in the table below, each included a number of sub-categories, and the number of targets fitting each of these sub-categories were pooled together under the more generic category shown in the table below.

category	Male Test	Female test	Male control	Female control
people	36	35	43	46
indoors	47	34	40	37
food	15	10	14	11
clothing	4	8	4	15
landscape	1	2	4	2
flora	25	10	28	8
fauna	20	10	17	8
buildings	13	11	14	2
vehicles	8	10	17	6
sports	6	7	8	7
other	11	6	4	1
water	18	9	20	7

The results of this analysis showed a normal distribution in the selection of targets for the male test and control targets, but a non-normal or non-parametric distribution in the selection of the female test and control targets. This led to a significant difference between the female and male control target selection categories, and a borderline significant difference between the female and male test picture selection categories, according to the category proportions. There was more variability in the target selection for the male, and less variability in the target selection for the female. However, when looking at the analytical results of the picture selection, (see table below) it was not clear how this would seriously affect the outcome of the experiment. It may be seen that the standard deviation of the test and control male target categories are similar, but there is a small difference between the standard deviation in female test and control target categories. However, it is difficult to see how this might seriously affect or skew the analysis of the results.

Summary statistics:

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
MT	12	0	12	1.000	47.000	17.000	13.545
FT	12	0	12	2.000	35.000	12.667	10.491
MC	12	0	12	4.000	43.000	17.750	13.261
FC	12	0	12	1.000	46.000	12.500	14.229

In analysing the target selection, statistical analysis of the target test and control categories for each subject was undertaken, to determine if there were significant differences which might have influenced the outcome of the experiment. However, analysis of the target selection data showed no significant difference between the target test and control categories for the female or for the male.

The statistical analysis comparing the female test target categories with the female control target categories is shown below followed by a similar comparison of the male test and male control target categories. From these analyses, it could be concluded that there is probably no significant effect of target selection bias in this experiment, which would skew the results of the overall data analysis.

It should be noted that non-parametric tests were used on the female target selection criteria data, as these showed a non-normal distribution. However normal parametric Ttest was used for the male target selection criteria data, as these showed a normal distribution.

PAIRED TTEST COMPARISON OF FEMALE TEST PICTURE CATEGORIES AND FEMALE CONTROL PICTURE CATEGORIES. (see below)

Summary statistics:

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
FT	12	0	12	2.000	35.000	12.667	10.491
FC	12	0	12	1.000	46.000	12.500	14.229

Sign test / Two-tailed test:

N+	6
Expected value	5.000
Variance (N+)	2.500
p-value (Two-tailed)	0.754
alpha	0.05

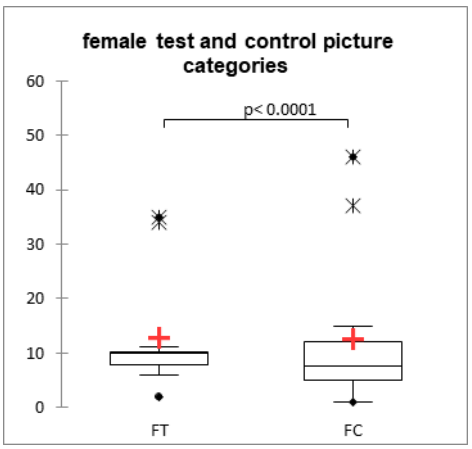
The p-value is computed using an exact method.

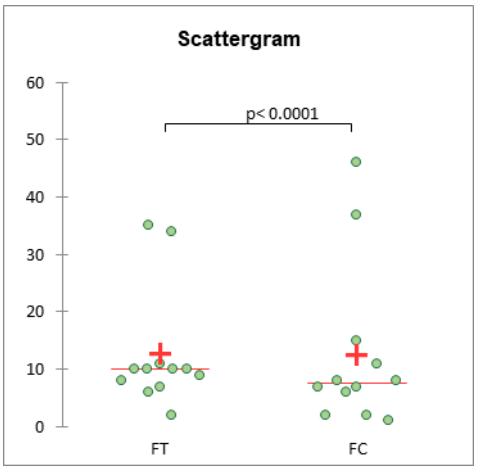
Test interpretation:

H0: The two samples follow the same distribution.

Ha: The distributions of the two samples are different.

As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.





PAIRED TTEST COMPARISON OF MALE TEST PICTURE CATEGORIES AND MALE CONTROL PICTURE CATEGORIES (see below)

Summary statistics:

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
MT	12	0	12	1.000	47.000	17.000	13.545
MC	12	0	12	4.000	43.000	17.750	13.261

t-test for two paired samples / Two-tailed test:

95% confidence interval on the difference between the means:

[-3.829, 2.329]

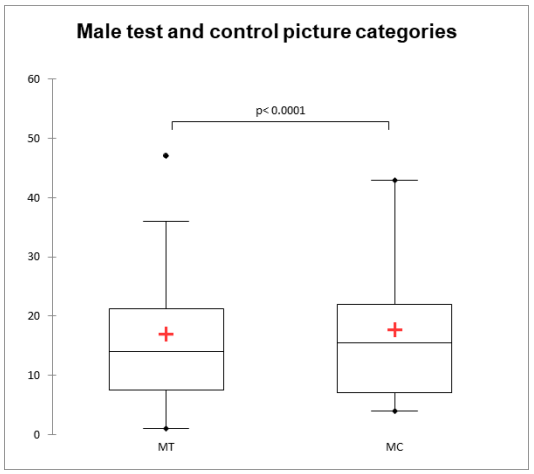
Difference	-0.750
t (Observed value)	-0.536
|t| (Critical value)	2.201
DF	11
p-value (Two-tailed)	0.602
alpha	0.05

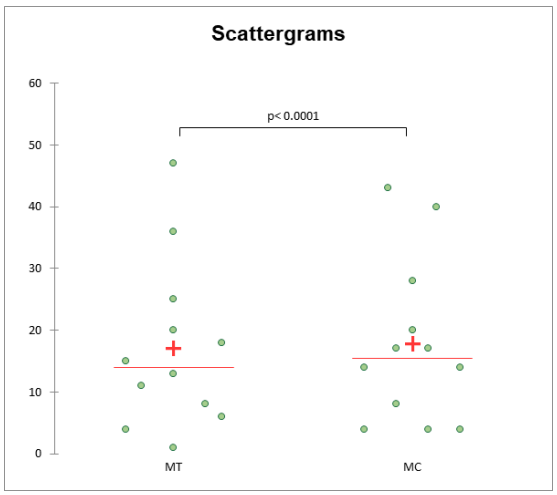
The number of degrees of freedom is approximated by the Welch-Satterthwaite formula

Test interpretation:

H0: The difference between the means is equal to 0.

Ha: The difference between the means is different from 0.
As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.





In analysing the results, great care was taken to guard against bias in scoring, by using an independent and blinded scorer.

However, it is possible that improvements to the scoring sheet could have been made, to make it easier for the scorer to match any of the observations in the subject's record appropriately. Some of the scoring sheet criteria were not easy to interpret, or could be misunderstood, for example, the distinction between the terms 'description' and 'perception' could be problematic.

There were also several instances when the scorer became aware of clues in the picture which helped him identify it as a control rather than test. This was due to items in the picture that were date dependent, and were later in date than the record. The scorer decided to omit these experiments from his scoring, due to possible bias in scoring, and no longer being blinded. This decision was upheld by the experimenter, and the few samples where suspected identification of control and test were made by the scorer, were excluded from the experimental analysis .

There were also several instances noted in subsequent quality control, where the scorer had failed to score a recorded item that the experimenter felt strongly matched a criteria, and a few where a difference of opinion might have led to different scoring. However, the independent scores of the scorer were accepted without any changes. Independence and blinding in scoring was considered essential. Overall there was consistency and discernment in the independent scoring, that fulfilled the requirements of the experiment. The judgement of the scorer was accepted by the experimenter as good and unbiased by any expectations. The scoring was accepted in full. The scorer had no idea who the subjects were, how many male or female. In fact he was surprised, subsequently, to learn that only two people were involved, and that the experimenter herself, who designed and coordinated the scoring, was also one of the subjects.

This was not a prospective experiment, in the sense that the score sheet and method of assessing or quantifying the data had not been formulated until near the end of the experimental period. Although all the test results had been dated and retained for future analysis, the method of analysis and strict conditions for each experiment were not formulated specifically in advance. This was an explorative experiment, conducted as an aid to development, rather than a strict scientific experiment. Therefore, from the point of view of scientific rigour, this is not an ideal experiment. However, in its favour, is the fact that this could be viewed as an explorative test of psychic perceptions, often categorized by mediums and psychics as clairvoyance, clairsentience,

claircognisance and clairaudience, and other extrasensory perception abilities in the community, in an ordinary setting, by ordinary, relatively untrained people, using methods they felt to be conducive, and testing those methods. In this experiment, environmental factors, such as lunar phase, geomagnetic field activity, atmospheric pressure, humidity and temperature, might also be correlated with results, at a later date. The opportunity to conduct experiments with the same two subjects over 6 years has allowed for a reasonable sample size for analysis. At the very least this experiment provides a decent sample size for statistical analysis and for further reflections about experimental protocol appropriate for such testing in future.

Unfortunately detailed description of the individual methods used to induce a 'psi-conducive' or receptive state were not always consistently noted and recorded by the subjects during the testing. However, both subjects were involved regularly, in a specific relaxation and spiritual attunement method several times a week, together, which induced a light trance state of altered consciousness. This altered state was described in the 'Mercury Light' blog, posted weekly on the internet, during this experiment, by the male subject, and confirmed, subsequently by a controlled experiment, analysing the electroencephelographic (EEG) recording of brain wave frequencies and amplitude proportions in both subjects over a number of weeks. This resulted in a good sample size of 46 experiments in all, and 46 matching controls. The results were highly significant, by statistical analysis, for both subjects, suggesting that both subjects knew how to regulate their brain waves to obtain an altered state of consciousness, which could be verified, in the case of the female subject, to coincide with mediumistic reception of detailed information purporting to come from deceased persons, and psychic perception related to living persons and events (see EEG report).

CONCLUSION

The results of the experiment reported here, suggest that extrasensory perception is probable in the case of the female subject, and possible, on occasion, in the case of the male subject. This means that there is a probability that extrasensory perception is taking place in this experiment.