Journal of Scientific Exploration, Vol. 12, No. 4, pp. 583–603, 1998 0892-3310/98

Experimental Studies on Telepathic Group Communication of Emotions

JAN DALKVIST AND JOAKIM WESTERLUND

Department of Psychology, Stockholm University
S-106 91, Stockholm, Sweden

Abstract — Four experiments testing for telepathic communication of emotions, evoked by slides, between a group of "senders" and a group of "receivers" using a total of 210 undergraduate students as subjects, are reported. Analysis of variance of aggregated data showed a significant difference among: a) believers in telepathy ($n = 53$), scoring below chance level, b) undecided ($n = 105$), scoring above chance level, and c) non-believers ($n = 48$), who did not deviate from chance level. Except for a minor deviation in the first experiment, the same pattern was obtained in all four experiments. The first picture presented exhibited a significant positive deviation from the value expected by chance. Hit score showed an inverted U-formed relationship to rated involvement, with negative scores for very high involvement values. Scores on an "openness" test, administered at the end of Experiment 4, did not show any relationship to performance, except for a significant negative correlation between a belief-in-telepathy item and hit score. In general, the results from the first three experiments were replicated in the fourth experiment, where all conceivable experimental errors had been eliminated. A so-called Monte Carlo Method, which is free from all statistical assumptions, confirmed the results obtained by traditional methods. Except for sampling error, no probable "natural" explanation of the findings could be found.

Keywords: parapsychology — ESP — telepathy — *psi* — belief — emotion

Introduction

Extrasensory perception (ESP) is a concept in parapsychology comprising various hypothetical perceptual responses to the outside world without any known sensory contact. Among these hypothetical phenomena, telepathy is the least questioned one (probably because it might be understood in analogy with transmission of information *via* radio or TV). Besides telepathy, ESP phenomena include clairvoyance (perception of objects or events that are not stimulating any known senses) and precognition (perception of future events that could not be anticipated on the basis of any known inferential process). Unfortunately, a telepathic experiment can not directly be discriminated from an experiment on clairvoyance or precognition, because the "receiver" may, conceivably, "see" the target directly (clairvoyance) or know beforehand what will happen in the experiment (precognition) without any help from the "sender."

The design of the four telepathy experiments reported in this paper was for

the most part developed some 20 years ago by one of us (JD), in cooperation with students in teaching a university course in experimental psychology. The main features of the design are as follows. In an attempt to increase the signal-to-noise ratio, not only one but several simultaneous senders were used. (Theoretically, the use of several senders might of course interfere with, rather than facilitate, the telepathic performance.) Moreover, only emotionally loaded stimuli were used, on the assumption that emotional objects or events are more easily transmitted telepathically than neutral ones due to their potential adaptive value during the course of biological evolution. Subjects were assigned to two groups. One group — the senders — were sequestered in a soundproof room and presented with a randomly ordered sequence of slide pictures, half of which were positively emotionally loaded and half of which were negatively emotionally loaded. The other group — the receivers — were isolated in another soundproof room. The receivers' task was to guess individually whether the picture presented to the senders in a given trial was positive or negative and to mark the guess on a form.

Peoples' belief in ESP has been a major variable in *psi* research. In her pioneering work, Schmeidler (1952; Schmeidler & McConnell, 1958) compared two groups of subjects in ESP performance, one called "sheep," who believed in ESP and the other "goats," who did not. The results confirmed what one might expect if ESP indeed exists: that the sheep on the average outperformed the goats. This finding, known as the "sheep-goat effect," is commonly regarded as one of the most consistent findings in parapsychological work (for a review, see Palmer, 1971), and is often taken as evidence supporting the authenticity of ESP. In the present series of experiments, belief in telepathy is one of the major independent variables; it is also the only one that was investigated consistently from the start of the project.

In an initial analysis of data from Experiments 1–3 (Dalkvist, 1994), nonparametric analysis of variance of aggregated data from the three experiments showed a significant ($p = 0.04$) overall level difference among (a) believers in telepathy ($n = 34$), scoring below chance level, (b) undecided subjects ($n = 38$), scoring above chance level, and (c) non-believers ($n = 24$), who did not deviate from the performance level expected by chance. Except for chance fluctuations, no probable "natural" explanation of the findings could be found.

Unfortunately, in Experiments 1–3, there was a possible, although unlikely, perceptual "leakage": the duration of a lamp signal informing the receivers when a picture was presented to the senders. Unlike the feeding of the slides, the lamp signal was not managed automatically in these experiments, but manually by one of the experimenters. By counting silently, the experimenter attempted to let the lamp shine for the first five out of the 20 seconds that the picture was presented. The possibility that there was a correlation between the duration of the signal, on the one hand, and type of picture shown, on the other, can therefore not be completely ruled out. In spite of the fact that, in Experiments 2 and 3, the experimenter who managed the signal lamp could

not see the pictures, it may be argued that the experimenter heard reactions from the subjects, but such reactions rarely occurred. Also, even if a correlation between the signal duration and type of picture did exist, it seems unlikely that this correlation could have caused the results obtained. First, use of signal duration information would require that receivers could discriminate and categorize durations of signals differing by tenths of a second and separated by relatively long intervals (more than 15 seconds). Second, it is difficult to see why signal length would be interpreted differently by subjects with different attitudes toward telepathy. Finally, only a minority of the receivers (say, one fourth) could see the signal lamp. (Unfortunately, these subjects can not be identified.) However, if this explanation is true, it would in itself be of considerable psychological interest.

In Experiment 4, the duration and timing of the lamp signal were controlled automatically. A minor extension was also made in Experiment 4. Assuming tentatively that telepathy does exist and that the different results obtained for groups of subjects differing in belief in telepathy in Experiments 1–3 reflect genuine telepathic phenomena, we wondered whether the different results were due to the attitude *per se,* or to some more general personality trait of which attitude toward telepathy is merely a reflection. It seems reasonable to assume that such a trait might be characterized by some form of "openness." In order to test that hypothesis, a questionnaire aimed to measure openness for alternative views was administrated after the main experiment.

In the above-mentioned report of Experiments 1–3 (Dalkvist, 1994), hit scores were only analyzed as a function of attitude toward telepathy. More recently, hit performance in these experiments, as well as in Experiment 4, have also been analyzed (a) as a function of psychological features of the stimulus pictures and (b) as a function of the position of a picture in the presentation sequence, independent of any particular property of the picture.

In the present paper, the four experiments considered above are presented together. The results from Experiment 4 will be compared with the merged results from Experiments 1–3. Moreover, the merged results from all four experiments will be considered.

Four students, all of whom had a skeptical attitude toward the existence of telepathy, served as experimenters in the first three experiments. Four new experimenters were used in the fourth experiment. These experimenters were graduate students who were paid for their work. They were all more or less skeptical about the existence of telepathy, and except for one of them — JW — they were not personally involved in the experiment.

Methods

Design

Each experiment comprised a number of sessions. In a given session, each

subject participated both as a sender and as a receiver, in order to obtain a maximal amount of data from the available subjects. The subjects were randomly assigned to two groups of equal or about equal size, one where they started as senders and one where they started as receivers. The number of sessions were 2, 3, 2 and 8 in Experiments 1 through 4, respectively, and the mean number of subjects per session was 9.5, 17, 18 and 13.9.

Subjects

Subjects were undergraduate psychology students. Table 1 gives a description of the subjects with respect to age, gender distribution and distribution of attitudes toward telepathy for each experiment separately and for the total set of data.

Before the subjects chose to participate in the experiment, they were informed that the experiment tested for telepathy and that strongly repulsive pictures, depicting, for example, criminal and war victims, were to be shown. It was recommended that students who knew they were very sensitive to such pictures not participate in the study.

The number of subjects and sessions was preset in each experiment, but in each experiment, some of the subjects who had agreed to participate did not show up.

Stimuli

Thirty slide pictures, fifteen with positive motifs and fifteen with negative motifs, described in the Appendix, were presented in each round. The same stimuli were used in all four experiments, except that two positive pictures were replaced with new ones in Experiment 4. As indicated in the Appendix, one of the two substituted pictures depicted an empty stroller, which at least one subject (according to a spontaneous remark in the response form) had regarded as negative, and the other depicted the moon in the sky, which could, conceivably, evoke a negative feeling.

In order to obtain a psychological description of the pictures being used,

TABLE 1
Characteristics of Subject Groups

	Age		Frequency Distributions					
			Attitude to Telepathy			Gender		
Exp. No.	Mean	Range	Non-believer	Un-decided	Believer	Females	Males	Total
1	32.0	20–44	5	7	7	16	3	19
2	27.1	20–54	14	18	18	29	22	51[a]
3	25.8	19–50	5	13	9	20	7	27
4	27.7	19–54	24	67	19	68	43	11[b]
1-4	26.7	18–50	48	105	53	133	75	208[b]

[a] One subject did not belong to any belief group.
[b] Two subjects did not belong to any belief group.

they were rated on four graphic (visual analog) scales, 100 mm in length. The ratings were made by 24 subjects, 19 females and 5 males, who had not participated in any of the main experiments. The scales were defined to measure how (a) pleasurable, (b) involving, (c) familiar, and (d) perceptible the pictures were.

The means and standard deviations of the four scales are given in Table 2. It should be particularly noted that positive and negative pictures are reliably separated on the pleasure scale, although the negative pictures are perceived as more displeasing than the positive pictures are perceived as pleasurable.

The correlations among the four scales are given in Table 3. As can be seen, most of the correlations are high, and all of them are significant. It may be noted that the familiarity, perceptibility, and pleasure scales are positively correlated, whereas the involvement scale is negatively correlated with the pleasure and familiarity scales.

Procedure. The senders and the receivers were sequestered in two acoustically insulated rooms, with one room between. The distance between the centers of the two experimental rooms was 7.90 m. The location of both of the rooms was evident to all subjects, since two experimenters received the subjects outside the doors of the two rooms.

The sound reduction from the sender room to the receiver room as well as the background sound in the receiver room were measured using a standard equipment. Employing pink noise as test sound, a reduction value of 41 db

TABLE 2
Means and Standard Deviations of Psychological Scale Values for Positive and Negative Target Pictures

	Type of picture			
	Positive ($n = 17$)		Negative ($n = 15$)	
	Mean (mm)	St. dev. (mm)	Mean (mm)	St. dev. (mm)
Pleasure	61.09	9.87	19.37	10.28
Involvement	52.91	8.51	69.47	7.95
Perceptibility	90.85	5.12	62.40	16.53
Familiarity	80.59	11.09	37.10	22.82

TABLE 3
Correlations Among the Four Rated Variables Characterizing Stimuli

	Pleasure	Involvement	Perceptibility	Familiarity
Pleasure	–			
Involvement	–0.73[c]	–		
Perceptibility	0.70[c]	–0.39[a]	–	
Familiarity	0.80[c]	–0.48[b]	0.89[c]	–

[a] $p < 0.05$
[b] $p < 0.01$
[c] $p < 0.001$

(50 dbA) was obtained. The background sound in the receiver room was measured to be 50 db (25 dbA). These measures indicate that weak normal sounds, such as sounds from breathing or bodily movements, would either have been totally absorbed or would have been impossible to distinguish from the background sound by the receivers, either consciously or subliminally.

As mentioned before, the sender and receiver rooms were connected by a signal device: a lamp in the receiver room that could be turned on and off from the sender room. As also mentioned before, the signal lamp was managed manually in Experiments 1–3 and automatically in Experiment 4.

The subjects received a booklet containing (a) a questionnaire, (b) written instructions, along with a description of the experimental situation, and (c) a response form for the experiment.

In Experiment 1, the questionnaire asked the subjects to indicate their age and gender and answer two questions: "Do you believe in telepathy?" and "Do you believe that you possess telepathic powers?" For each question, the subject was to choose one of the three alternatives: "yes," "no," and "don't know." The last question was ignored, however, because only one subject answered affirmatively to that question. In Experiments 2–4, only the first question was asked.

In addition to the written instructions, the senders were given oral instructions, emphasizing the information specific to the senders. They were instructed to concentrate on the positive or negative feeling evoked by each shown picture and to hold that feeling as long as the slide was presented; they were thus not instructed to convey the feeling to the receivers. The senders were also told that they could shut their eyes or leave the room if a picture was experienced as unbearably unpleasant. They were told to be silent.

There were two experimenters in the sender room. In Experiments 1–3, one of the experimenters presented the slides and provided the oral instructions, and the other experimenter controlled the signal device. In Experiment 4, where, as mentioned, the signal device was run automatically, one of the experimenters provided the oral instructions, and the other experimenter presented the slides. The two experimenters and the projector were located behind the subjects. In Experiments 2–3, the experimenter who handled the signal lamp sat with his back turned to the wall on which the slides were projected so that he could not see them; this was done in order to reduce the risk that the length of the lamp signal would be affected by the picture being presented. In Experiment 4, both experimenters were turned away from the projection wall, in order to prevent them from becoming "additional senders."

The slides were presented in random order, one order for each run, except that the same two random orders were used in both sessions of Experiment 1. In Experiment 1, the random orders were obtained using a random number table. In Experiments 2–4, the random orders were obtained using a computerized random number generator (in the program library LOTUS 123, version 2.01). Each slide was presented for 20 seconds, with an inter-stimulus interval

of about half a second (unfortunately, in Dalkvist (1994), the interval was erroneously indicated to be about one second). The signal lamp in the receiver room was turned on immediately when a picture was inserted into the projector and turned off after about five seconds.

Like the senders, the receivers were given oral instructions emphasizing the information specific to them. The receivers were instructed to imagine the presentation of pictures in the sending room. They were informed about the number of slides and that the pictures were chosen to evoke either positive or negative feelings (but not that the number of positive and negative pictures were equal). The receivers were also told that one of the experimenters would tell them each time the senders were presented with a new picture, including the first one, and how this was done. The receivers were instructed to guess whether a given picture was positive or negative and to mark the guess with a cross on the form; they were forced to choose one of the two alternatives.

In order to prevent them from seeing each others' response forms, the receivers were seated in a circle with their backs turned to each other. There were two experimenters in the room. One of them watched the signal and told the subjects when the lamp was lit, by announcing the order of each picture (Picture 1, Picture 2...). The other experimenter provided the oral instructions.

In the middle of the session, the subjects changed rooms. Those who started as senders now served as receivers and *vice versa.* (In one session in Experiment 3, the second part of the experiment was not completed, because one [female] subject fainted when serving as sender.)

At the end of each session in Experiment 4, subjects received a questionnaire containing 17 statements selected from an openness questionnaire, developed by Westerlund (1983), openness being defined in terms of nonextremeness in attitude. Subjects were instructed to respond to each statement by checking one of seven positions on a scale ranging from "agree" to "do not agree." A response at each end of the scale (positions 1–2 or 6–7) was interpreted as low openness to different attitudes and was coded as "0," whereas a response in the middle of the scale (positions 3–5) was interpreted as high openness to different attitudes and was coded as "1." The statements concerned various subject matters, but one of them happened to be related to belief in telepathy: "Some people have telepathic powers."

Immediately after each session in Experiment 4, the data were put into envelopes, which were sealed and signed by the four experimenters. When the experiment was finished, the envelopes were broken, and the material was copied for practical use. The original data were then saved in sealed and signed envelopes. This procedure was supervised by a notary public (Björn Edlund, lecturer at the Department of Psychology, Stockholm University, and not involved in the project).

Results

Attitudes

Table 4 shows the mean difference between observed and expected hits and the corresponding standard deviations for Experiments 1–3, Experiment 4, and Experiments 1–4, respectively.

As can be seen, the pattern of mean scores obtained in Experiment 4 is essentially the same as that obtained in the three previous experiments: the believers in telepathy provided a negative mean score, the undecided group a positive one, and the mean score for the non-believers fell between those for the two other groups, although the score obtained for the "no" group in Experiment 4 was higher than that obtained for the same group in the previous experiments. None of the three mean scores obtained in Experiment 4 differed significantly from zero, however, but the mean score for the undecided group differed significantly from that for the "yes" group ($t(36.47) = 2.11$; $p = 0.04$, two-tailed).

There is also a difference between Experiment 4 and the previous experiments with respect to the pattern of standard deviations: whereas the "no" group showed a significantly larger standard deviation than the two other groups — particularly the undecided group — in the previous experiments, there is no marked difference between the standard deviations in Experiment 4.

In the total data set (Experiments 1–4), a distinct pattern of mean hit scores appeared: both the positive score for the undecided group and the negative score for the "yes" group differ significantly from zero, whereas the score for the "no" group is close to zero.

Table 5 shows the results of analysis of variance (ANOVA) testing for overall score level differences among the three belief groups in Experiments 1–3, Experiment 4, and Experiments 1–4, respectively. As can be seen, whereas neither the analysis of the data for Experiments 1–3 nor the analysis of the

TABLE 4
Descriptive Statistics for the Difference Between Observed and Expected Hits and the Corresponding Standard Deviations for Experiments 1–3, Experiment 4, and Experiments 1–4, Respectively

	Experiments								
	1-3			4			1-4		
Belief in telepathy	*n*	Mean	St. dev.	*n*	Mean	St. dev.	*n*	Mean	St. dev.
"no"	24	–0.20	3.80	24	0.44	2.94	48	0.14	3.40
"?"	38	0.87[b]	2.15	67	0.51	2.86	105	0.64[b]	2.62
"yes"	34	–0.70	2.79	19	–0.90	2.41	53	–0.76[b]	2.64
Total	97[a]	0.06	2.91	111[b]	0.25	2.83	208	0.16	2.86

[a] One subject did not belong to any belief group.
[b] Deviates significantly from zero at the 5% level (two-tailed).

TABLE 5
Summary of Results from ANOVA for Differences Between Observed Hits and Expected Hits in the Three Belief Groups in Experiments 1–3, 4, and 1–4, Respectively

Experiments	F	df	p
1-3	2.81	2/93	0.07
4	1.87	2/107	0.16
1-4	4.40	2/205	0.01

data for Experiment 4 reached significance, the analysis of the total set of data did so, at the 1% level.

Position Effects

When plotted against position order in the sequence of picture presentations, hit rate for individual pictures did not show any general trend, such as an overall increasing or decreasing trend.

A more specific finding related to the position order of the pictures was made, however. Although no general declining trend was found, the first picture showed a strong tendency to exhibit a higher hit rate than the remaining pictures. Thus, the first position in the presentation order yielded the highest hit rate of all positions in the "yes" and undecided groups, and the next highest hit rate in the "no" group.

(In Dalkvist & Westerlund, 1995, hit rate was claimed to exhibit an inverse function of position order. Subsequent analyses showed, however, that the good fit of this function was altogether attributable to the first-picture effect.)

In the total group, the first position produced a hit rate of 0.62, which differs from chance at a high significance level ($\chi^2 = 13.0$; $p = 0.0003$). In Experiments 1–3, the hit rate for the first position was 0.59, which only approached significance ($\chi^2 = 2.98$; $p = 0.08$), and in Experiment 4 it was 0.66 ($\chi^2 = 11.04$; $p = 0.0009$).

Hit Rate as Related to Picture Characteristics

Of the four rating scales characterizing the stimulus pictures, the involvement scale showed the clearest indication of being related to hit rate. In the upper panel of Figure 1, the hit rate for the total group in Experiments 2–4 has been plotted against rated involvement. (Unfortunately, the data from Experiment 1 could not be used, because the order of presentation was not documented.) As can be seen, the data show an inverted U-formed trend, with a clear tendency for the hit rates to fall above baseline for moderately strong involvement. For still higher involvement levels, the hit rates fall below baseline. The curve represents a highly significant ($F(2/27) = 5.67$; $p = 0.009$) fitted second order polynomial function, as indicated by ANOVA. Significant regression weights were obtained for both the linear ($t = 2.78$; $p = 0.01$) and the quadratic component ($t = -2.94$; $p = 0.007$) of the function.

In order to test whether the relationship shown in Figure 1 could be attributed to the strong first picture effect considered in the previous section, the relationship between hit rate and involvement was examined with the first target eliminated from the data. The inverted U-trend remained unchanged.

The data for the three attitude groups in Experiments 2–4 have been plotted separately in the three small panels in Figure 1. Like the plot for the whole group, each of the plots for the three attitude groups exhibits an inverted U-formed trend, as indicated by the fitted second order curves. Only the fit for the "no" group is significant ($F(2/27) = 5.70$; $p = 0.009$), however. As can be seen from the graphs in Figure 1, the positive pictures (open circles) are characterized by relatively low involvement values, whereas the negative pictures (filled circles) are characterized by high involvement values. This means that hit rate for the positive pictures is positively related to rated involvement, whereas hit rate for the negative pictures is negatively related. It is noteworthy that negative pictures with very high involvement values show a strong tendency to fall below baseline, that is, to exhibit negative hit rates.

Table 6 shows the linear correlations between hit rate and involvement for positive and negative pictures, respectively, in the three attitude groups and the total group. As could be expected from inspection of Figure 1, each set of data exhibits a positive correlation for the positive pictures and a negative cor-

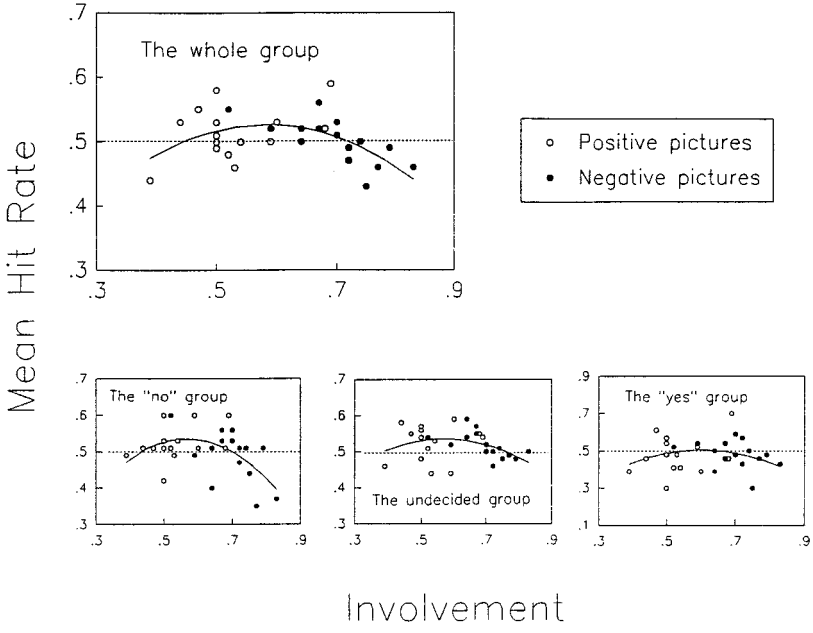

Fig. 1. Hit score plotted against rated involvement in response to the thirty stimulus pictures for the total subject group (Experiment 1–4) and the three belief groups in that group. Empty circles denote positive pictures and filled circles negative pictures. A second order polynomial function was fitted to each plot.

TABLE 6
Correlations Between Rated Involvement and Hit Rate for Respective Positive and Negative Pictures in Experiments 2–4

Belief in telepathy	Type of picture	
	Positive	Negative
No	0.37	–0.65[a]
Undecided	0.08	–0.60[a]
Yes	0.19	–0.32
Total	0.33	–0.71[b]

[a] $p < 0.05$
[b] $p < 0.01$

relation for the negative pictures. The correlations are stronger for the negative pictures, however. It is noteworthy that the correlation obtained for the negative pictures in the total group, –0.71, explains as much as 50% of the variance.

As could be expected from the correlations among the four rating scales, shown in Table 3, the involvement scale was not the only scale that was related to hit rate, even though none of the relationships for the three remaining scales was found to be significant. Of the three scales, the familiarity scale showed the clearest tendency to be related to performance, in spite of the fact that the pleasure scale was more strongly correlated with rated involvement than was the familiarity scale. Thus, in agreement with the negative correlation between the involvement and the familiarity scales, together with the results shown in Figure 1, hit rate showed a tendency to be related to the familiarity scale in a U-formed fashion for each of the three belief groups as well as the total group in Experiments 2–4.

Openness as Related to Hit Scores

The openness questionnaire showed a moderately high reliability (Cronbach's alpha = 0.46).

The sum of each subject's openness scores for the 17 statements was used to measure general openness. This measure did not correlate significantly with hit score ($r = –0.08$; $p > 0.10$), however.

We also examined whether actual attitude, as opposed to extremeness in attitude, was correlated with hit score for any of the 17 statements. Only attitude to the aforementioned statement concerned with telepathy ("Some people have telepathic powers") was significantly correlated with hit score ($r = 0.22$; $p = 0.02$). The more the subjects opposed the statement, the higher was his or her hit score.

Gender and Task Order Comparisons

Gender. Using the total set of data (Experiments 1–4), gender comparisons were made with respect to each of the major independent variables considered

above: attitude toward telepathy as measured before the experiment, position order, and rated involvement. A gender comparison was also made with respect to attitude to telepathy as measured after the experiment in Experiment 4.

As can be seen from Figure 2, both males and females showed the same relationship between hit score and attitude toward telepathy (as measured before the experiment) as was found for the mixed gender groups considered above: subjects with an undecided attitude showed a positive mean score, those believing in telepathy a negative mean score, and the skeptics a mean score close to zero. Accordingly, a two-way ANOVA, with gender and attitude toward telepathy as independent variables, failed to reveal a significant gender effect ($F(1/205) = 0.09$; $p = 0.76$) or a significant gender *vs.* attitude-toward-telepathy effect ($F(2/205) = 0.02$; $p = 0.98$). However, whereas the effect for attitude toward telepathy was significant for the females, as indicated by a one-way ANOVA ($F(2/129) = 3.65$; $p = 0.03$), it was far from significant for the males ($F(2/71) = 0.85$; $p = 0.43$).

As to the above-mentioned correlation between attitude toward telepathy as measured after the experiment in Experiment 4 and hit score, only the females were responsible for that correlation. Thus, whereas the females ($n = 68$) showed a highly significant correlation ($r = 0.33$; $p = 0.006$), the correlation obtained for the males ($n = 43$) approached zero ($r = 0.08$; $p = 0.62$).

The inverted U-formed relationship between hit rate and rated involvement invariably obtained in the mixed gender group was also obtained for males and females separately. As can be seen in Table 7, the fit of a second order function was equally good for males and females, although neither the curve for the

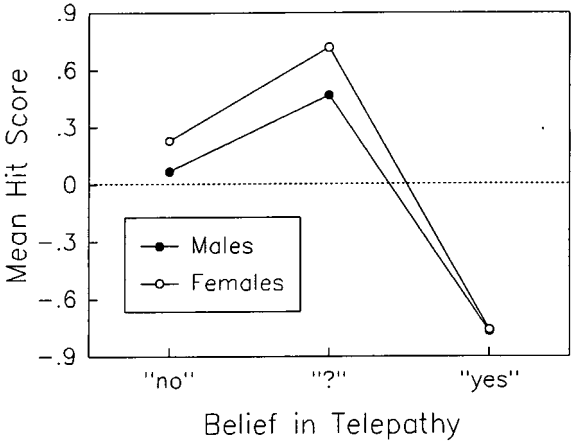

Fig. 2. Hit score plotted against belief in telepathy as measured before the experiment for males (filled circles) and females (open circles).

TABLE 7
Summary of Results from ANOVA Corresponding to Fits of a Second-Order Function Relating Hit Score to Rated Involvement in Response to the Pictures for Males and Females

Gender	*F*	*df*	*p*
Male	2.26	2/27	0.12
Female	2.56	2/27	0.10

males nor the curve for the females reached significance. As to the first-picture effect, there was no significant difference between males and females.

Task Order. In order to investigate whether the order of the two tasks — sending a message and receiving a message — had any effect on the results, the subjects were divided into two groups, one comprising the subjects who started as receivers and the other the subjects who started as senders.

As can be seen in Figure 3, both groups exhibited the same relation between hit rate and attitude toward telepathy as was obtained in previous analyses, although the pattern was somewhat more pronounced for the subjects who started as senders. Accordingly, a two-way ANOVA did not show either a significant task order effect ($F(1/205) = 0.30$; $p = 0.58$) or a significant task order *vs.* attitude-toward-telepathy effect ($F(2/205) = 0.46$; $p = 0.63$). Furthermore, although the overall attitude effect was significant, as noted before, a one-way ANOVA failed to reveal a significant attitude-toward-telepathy effect, both for the subjects who started as receivers ($F(2/104) = 1.84$; $p = 0.16$) and for those who started as senders ($F(2/96) = 2.66$; $p = 0.08$).

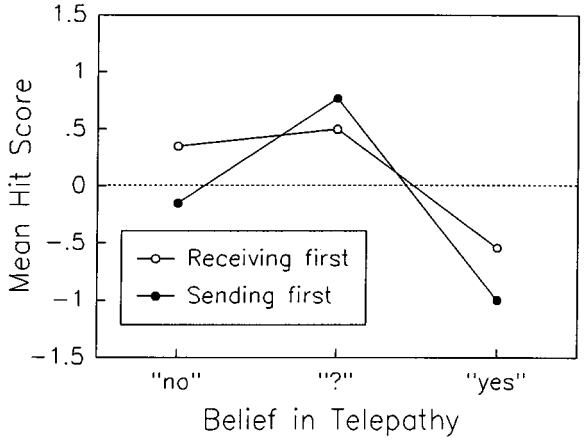

Fig. 3. Hit score plotted against belief in telepathy as measured before the experiment for subjects who started as senders (filled circles) and subjects who started as receivers (open circles).

Test of Random Orders

For each of the 25 different random orders used in the above studies, the occurrence of a positive picture was coded as "1" and the occurrence of a negative picture as "0."

A so-called run test (a non-parametric test which is specifically developed in order to test for deviations from a random order in a series of binary values) failed to show any evidence that the orders were not genuine random orders; thus, in accordance with the null hypothesis, only one of the 25 series exhibited a z-value deviating significantly from zero at the 5% level, and no series showed a higher absolute z-value.

A further test of the random orders being used were done using the Monte Carlo method, described below.

The possible occurrence of a correlation between the two random orders used within a given session, which might, conceivably, have been used by the subjects who started as senders in their subsequent role as receivers, was also tested. None of the 12 correlations obtained reached significance, the mean correlation being –0.08, with a standard deviation of 0.13.

Test of Statistical Assumptions

Because they are parametric, all the ANOVAs reported above require that the assumptions of equal variances, normal distributions, and statistically independent measures are satisfied. Likewise, because the correlations reported so far are all product-moment correlations, they require the assumptions of normal distributions and independent measures to be fulfilled. In the case of χ^2, used to test the first-picture effect, only the assumption of independent responses has to be satisfied, because it belongs to the class of nonparametric methods.

Of the three assumptions mentioned in the previous paragraph, that of independent measures is the most problematic one in the present experiments for two reasons. First, the use of more than one receiver at a time could have created statistically dependent responses. For example, the subjects in the same receiver group might, consciously or unconsciously, have influenced each other, or have been influenced collectively by the experimenters in the receiver room. Another possible source of statistical dependence in the present experiments is that all or some of the subjects in a given receiver group exhibited a common response bias, such as generating more positive or negative responses at the beginning of the experiment than at the end of it. The second reason why the assumption of independent measures is particularly problematic is that it can only be tested indirectly.

Effects of violations of one or more of the three assumptions discussed above have been tested by means of the Monte Carlo method — an inferential computer simulation method which does not require any statistical assumption at all to be satisfied. In such a method, the p-values for a given statistical para-

meter, say the *F*-ratio for a certain effect in ANOVA, is not obtained from a corresponding theoretical sampling distribution, as in standard methods, but from a corresponding simulated "empirical" sampling distribution. In the present version of the Monte Carlo method, the sampling distribution is generated by randomly exchanging the random order of a given target sequence for the random order of another target sequence occurring in any of the four experiments a large number of times — 1000 times or more in the present tests — and computing the relevant statistical parameter for each new random order.

The present version of the Monte Carlo method also provides a test of effects of a possible common bias in the random orders being used. While such an effect may lead to spuriously significant results using conventional statistical tests, the effect does not produce any significant results at all using the present Monte Carlo method. As will be developed in a paper in progress, the reason is that in the present Monte Carlo method, the effect of a given random order bias is involved not only in the original calculation of the statistical parameter being tested, but also in all the simulated calculations of that parameter.

Except for the correlation and regression analyses of hit rate as related to picture characteristics, where the assumption of independent responses did not offer any problem because pictures, and not individuals, were compared, all the above parametric analyses showing positive results were compared with corresponding Monte Carlo analyses. The agreement obtained was surprisingly good. As a matter of fact, the *p*-values obtained using the Monte Carlo method was slightly smaller than the *p*-values obtained using a conventional method, suggesting that occurring violations of assumptions tended to decrease rather than increase the statistical power.

A good agreement was also obtained between the original χ^2 test of the first-picture effect for the total set of data and a corresponding Monte Carlo test. In this case, however, the *p*-value obtained by the Monte Carlo method ($p = 0.001$) was somewhat larger than that obtained using the standard method ($p = 0.0003$), suggesting that the first picture was somewhat impaired by a common response bias.

To summarize, our Monte Carlo tests indicated that effects of violations of statistical assumptions were small and tended to decrease rather than increase the statistical power. In particular, the use of more than one receiver at a time did not create any serious statistical bias.

Discussion

Attitude Effects

The three attitude groups showed the same pattern of mean hit scores in Experiment 4 as in the previous experiments: the undecided group exhibited positive scores, the "yes" group negative scores, and the scores for the "no" group fell between those for the other two groups. The results were not significant,

however, except for a significant difference between the mean scores for the "yes" group and the undecided group. The explanation may be that, for some reason, the distribution of attitudes in Experiment 4 differed somewhat from those in the previous studies, with a larger proportion of subjects in the undecided group and a smaller proportion in the "yes" group. In any case, by merging the data from Experiment 4 with the previous data, several clearly significant attitude effects could be demonstrated.

Further evidence for a relationship between attitude toward telepathy and telepathic performance was provided by the fact that the only item in the openness questionnaire that showed a significant relationship with hit score was the "belief-in-telepathy" item. The finding of a purely linear relation between that item and hit score rather than a curvilinear relationship, which would be expected from the pre-experiment attitude results, is interesting. The discrepancy reflects the fact that some subjects rated their attitude to telepathy differently before and after the main experiment. This may be due to different measurement methods. It is also conceivable, however, that some subjects changed their attitudes after their experiences during the experiment.

The present relationships found between attitude toward telepathy and telepathic performance are inconsistent with findings of previous studies addressing the relationship between ESP performance and attitude toward ESP, referred to in the introduction. Thus, in 13 out of 17 studies discussed by Palmer (1971), the "sheep" (those subjects who believed in ESP) outperformed the "goats" (the non-believers). However, only three of the 17 studies tested exclusively for telepathy, as opposed to clairvoyance or precognition. The results from one of these studies (Nash, 1965), testing for telepathic communication of digits between a "sheep" and a "goat," are consistent with those from the present series of studies in that they demonstrate negative performance scores for subjects who believed in telepathy. However, the results of the two other studies (Moss, 1969; Moss & Gengerelli, 1968) which, like the present experiments, dealt with the telepathic response to a sender's emotional reactions to slides, contradict the present results; thus, in both studies, the "sheep" tended to perform better than the "goats." One of the many differences between those two experiments and the present experiments is that strongly repulsive pictures were not used in the former but were used in the latter experiments.

Position Effects

An apparent position effect in parapsychological experiments was observed long ago by Rhine, "the father of parapsychology," and others. Reportedly, performance tended to decline over trials, possibly with a rise at the end of the experiment (Rhine, 1971).

In the present series of experiments, no general decline effect was found. A similar, although more specific effect was obtained, however: a marked drop in hit rate from the first to the second picture in the series, with a significant

above-chance-level hit rate for the first picture. This apparent first-picture effect might be attributed to the fact that the first picture cannot be interfered with by any other picture in the series (provided, of course, that precognition is not involved.)

Effects of Picture Characteristics

Further evidence for telepathic information transmission was obtained by analyzing hit rate as a function of rated psychological characteristics of the stimulus pictures. Even though the four rating scales were strongly correlated with one another, only the involvement scale seemed to be basically related to hit rate. For the total set of data, a highly significant inverted U-formed relationship, with negative hit rate for the most involving pictures, was obtained. The fact that all three attitude groups showed the same relationship argues for considering the relationship as genuine, even though only the relation for the "no" group was significant. This interpretation is supported by the fact that the relationship appeared both in the merged data from Experiments 2 and 3 and in the data from Experiment 4, even though it was non-significant in the former case.

If the relationship in question is not attributable to sampling errors, but is real, how, then, could it be explained? That hit rate increased with increasing rated involvement, as it did for low and moderate involvement levels, seems reasonable; this could be explained in terms of a positive correlation between involvement and concentration. The fact that the hit rate declined when involvement increased above a certain level is unexpected, however, particularly since negative hit rates were obtained for pictures with very high involvement values.

A clue to understanding the hit rate curve for involvement may be the fact that the inclining part of the curve represents the positive pictures, whereas the declining part represents the negative pictures, showing a strong, negative linear correlation, with negative hit scores for the most involving pictures. One interpretation of the declining hit score for the negative pictures would be that the subjects exhibited some form of perceptual defense in response to the most involving negative pictures. However, the most involving negative pictures were not always rated as very displeasing; most notably, Picture 30 (oil-injured bird in a human hand) was rated as extremely involving but only as moderately displeasing. Thus, according to the perceptual defense interpretation, the defense reaction must have been in response to personally involving pictures rather than against displeasing and repulsive ones, such as Picture 6 (naked body partly in view, signs of severe assault) or Picture 10 (prostrate torture victim with wounds on his back).

An alternative interpretation is suggested by the fact that negative pictures characterized by high involvement and negative hit rates — such as the oil-injured bird picture just mentioned — show a strong tendency to evoke compassion (according to our own judgments). Perhaps there is a similarity between

compassion or sympathy, on the one hand, and positive feelings such as love and experiences of beauty, on the other hand, that tend to give rise to "signal reversals," which, in turn, result in negative hit rates for a subset of the negative pictures.

Needless to say, the above interpretations are quite speculative and should not be taken too seriously. However, they may be useful as guides for future research.

Gender Differences

There were no significant gender differences, and, in general, males and females showed the same pattern of results. However, the patterns appeared to be more pronounced for the females than for the males, and no significant results were obtained for males only. It would, of course, be of great interest if this suggested superiority of females in receiving telepathic messages could be replicated in future research.

General Conclusions

Taken together, the results from the four studies reported in this paper seem to lend strong support to the view that telepathy (or some other form of ESP) does exist and is possible to demonstrate in controlled experimental situations. This conclusion is based on the judgment that no reasonable alternative explanation of the results seems to exist.

Of the conceivable alternative explanations, the possibility that the results are attributable to chance factors is perhaps the least improbable one. The likelihood that this explanation is correct is small, however, even though no exact figure can be calculated. In a conventional psychological experiment, where the results never challenge the prevailing world view, the present results would never be suspected to have been caused by sampling errors.

The experimental controls were increasingly sharpened in the four experiments, and in Experiment 4, where automatic control of the signal light was introduced, no one has been able to detect any source of experimenter bias/intrusion or perceptual leakage. Thus, it seems that such explanations can be excluded.

It is important to emphasize that the findings could not have resulted from subjects having produced some particular, systematic sequence of responses. This would have required that the random orders had been biased and correlated with the response sequences, or were known beforehand by the subjects. Our random order tests and, particularly, our Monte Carlo tests indicate that the positive results cannot be due to systematic response sequences in combination with biased random orders.

As mentioned before, however, there is a theoretical possibility that the subjects who started as senders could have gained some information about the random orders. Thus, even if the random orders were perfectly unbiased, suc-

cessive random orders might have been correlated, and the subjects who started as senders might have learned part of the random order and used it in their subsequent role as receivers. In reality, however, there was no significant correlation between the two random orders in any session. Moreover, the results for the subjects who started as senders did not differ significantly from the results for those who started as receivers.

It should also be kept in mind that the number of subjects and sessions was preset in each experiment, thus excluding the possibility that the experiment was terminated when the results happened to be positive due to chance fluctuations.

Admittedly, all assumptions for the statistical tests were not satisfied. Perhaps most important, the intra-subject responses were certainly not uncorrelated. Although such violations of statistical assumptions can never give rise to positive results, they can affect the p-values obtained in the statistical tests, making them either smaller or larger than the p-values that would have been obtained if the assumptions had been satisfied. Fortunately, the Monte Carlo simulations indicated that occurring violations of statistical assumptions did not give rise to any significant result but, actually, tended to increase rather than decrease the p-values.

Remaining alternative explanations are errors in data handling and deliberate cheating. It is not realistic to think that data handling errors, such as erroneous entering of data into the computer, could account for all or most of the results obtained, particularly in view of the fact that large parts of the handling were cross-checked. It is understandable that suspicions of cheating tend to arise when positive parapsychological results are presented, particularly in view of the fact that cheating has in fact been revealed in parapsychological research. (It is fair to add that cheating has also been revealed in ordinary psychological and other scientific research.) As to the possibility of cheating in the present studies, in Experiment 4, we have tried to protect ourselves from suspicions or accusations of cheating by saving raw data in sealed and signed envelopes, making it possible for anyone to check our results. Moreover, the experimenters in Experiments 1–3 did not know those in Experiment 4, thus excluding the possibility of a collective experimenter fraud.

Convincing results in support of the existence of telepathy have also recently been obtained by Honorton and his co-workers (Honorton *et al.,* 1990; Bem & Honorton, 1994) using a so-called ganzfeld design, where the receivers are subjected to partial sensory deprivation, giving rise to the perception of a homogenous perceptual field, or a ganzfeld, in order to reduce the "noise" produced by ordinary perception. (The subjects are typically relaxing in a comfortable chair in a sound-proof room, with their eyes covered by ping-pong ball halves illuminated by homogeneous light and their ears with earphones emitting white noise.) Whereas previous ganzfeld studies have been sharply criticized for being poorly controlled and documented (Hyman, 1985), the recent series of studies can not be criticized on these grounds. Yet an overall analysis

of the 11 studies which enter in the series, reveals a mean hit rate of 32%, to be compared with a hit rate of 25% expected by chance. This hit rate is statistically significant, with a p-value of 0.002. Moreover, all the 11 experiments, except one, exhibited positive hit rates. It is encouraging that two different series of studies, using different designs, seem to lead to the same conclusion.

Acknowledgements

This paper is based on a report from the Department of Psychology, Stockholm University (Dalkvist & Westerlund, 1995). The revisions made are mainly the result of subsequent re-analyses.

The research reported in this paper was supported by grant from John Björkheim's memorial foundation.

We are indebted to Ingvar Bergman, Åsa Eriksson, Daniel Karlstedt and Bjarne Olsson for their valuable assistance in the first three experiments reported and to Eleonore Jansson, Maarit Johnsson and Lennart Sjöberg for their valuable assistance in the last experiment. We are also indebted to the unknown students who, long ago, conducted the experiment in which the present research originates.

Correspondence and requests for reprints should be addressed to: Jan Dalkvist, Ph.D., Department of Psychology, Stockholm University, S-106 91 Stockholm, Sweden.

References

Bem, D. J., & Honorton, C. (1994). Does *psi* exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, 115, 4.

Dalkvist, J. (1994). Telepathic group communication of emotions as a function of belief in telepathy. *Reports from the Department of Psychology*, Stockholm University, No. 785.

Dalkvist, J. & Westerlund, J. (1995). Four experiments on telepathic group communication of emotions. *Reports from the Department of Psychology,* Stockholm University, No. 798.

Honorton, C., Berger, R. E., Varvoglis, M. P., Quant, M., Derr, P., Schechter, E. I., and Ferrari, D. C. (1990). Experiments with an automated testing system and a comparison with a meta-analysis of earlier studies. *Journal of Parapsychology,* 54, 99.

Hyman, R. (1985). The ganzfeld *psi* experiment: A critical appraisal. *Journal of Parapsychology,* 49, 3.

Moss, T. (1969). ESP effects in "artists" contrasted with "non-artists." *Journal of Parapsychology,* 33, 57.

Moss, T., & Gengerelli, J. A. (1968). ESP effects generated by affective states. *Journal of Parapsychology,* 32, 90.

Nash, C. B. (1965). Opposite scoring direction in ESP. *Journal of Parapsychology,* 29, 115.

Palmer, J. (1971). Scoring in ESP tests as a function of belief in ESP. Part I: The sheep goat effect. *Journal of the American Society for Psychical Research,* 65, 373.

Rhine, J. B. (1971). *The Research of the Mind.* New York: William Sloane. (Originally published 1947).

Schmeidler, G. R. (1952). Personal values and ESP scores. *Journal of Abnormal and Social Psychology,* 47, 757.

Schmeidler, G. R., & McConnell, R. A. (1958). *ESP and Personality Patterns.* New Haven, CT: Yale University Press.

Westerlund, J. (1983). Intolerans mot tvetydighet — en entydig variabel? (Intolerance of ambiguity — an ambigous variable?) University of Stockholm, Department of Psychology, BC-uppsats nr. 37.

Appendix

Description of Stimulus Pictures

Picture No.	Type of picture	Picture description
1	Positive	Father taking picture of mother and two children in a stroller
2	Negative	Women with a knife stabbed into her chest
3 Exp.1–3	Positive	Moon in the sky
3 Exp. 4	Positive	Man in a mountain landscape
4	Negative	Man lying on the ground and another man running away with a machine gun in his hand
5	Positive	Coastal view with mountains
6	Negative	Naked body partly in view, signs of severe assault
7	Positive	Woman looking at a mountain landscape
8	Negative	Severely emaciated man
9	Positive	Little girl playing with a doll on the grass, woman watching
10	Negative	Prostrate torture victim with wounds on his back
11	Positive	Three women bathing nude (Anders Zorn painting)
12	Negative	Boy with skin injuries on his face and on his hands sitting on a bed looking at the viewer
13	Positive	Setting of the sun
14	Negative	Dead soldiers lying on biers and mourning comrades
15	Positive	Two female friends looking with interest at something, perhaps a piece of embroidery (Anna Anker painting)
16	Negative	Black boy with a severely demolished foot, half-lying on a stretcher, and another boy helping
17 Exp.1–3	Positive	Two empty strollers
17 Exp. 4	Positive	Baby's feet against adult hand
18	Negative	Traffic victim with crushed skull on the ground and a squatting man looking at him
19	Positive	A collection of teddy-bears
20	Negative	Dead man with a knife stabbed deeply into his chest
21	Positive	Butterfly sucking honey from a flower
22	Negative	Dead man with wounds on the back of his head and neck and a needle inserted into the neck
23	Positive	Face of a smiling elderly man
24	Negative	Wounded ear with blood
25	Positive	Two girls reading a book together with their heads close together (painting)
26	Negative	Emaciated children lying on the ground, one looking into the camera
27	Positive	Asiatic looking wedding couple
28	Negative	Dead man lying on the ground with a gun next to his head
29	Positive	Flowers and a bumble bee
30	Negative	Oil-injured bird held in a human hand