Journal of Parapsychology, Vol. 54, March 1990

A COOPERATION-COMPETITION PK EXPERIMENT WITH COMPUTERIZED HORSE RACES

By George P. Hansen

ABSTRACT: To test whether PK influence is additive, an experiment was conducted in which pairs of subjects participated in a computer-simulated horse race under conditions of cooperation and competition. The race was among four horses animated on an Apple computer. The winning horse was the first to accumulate 120 hits from binary decisions (p = .5) generated by a Bierman-RIPP RNG. For the game, each member of a pair of subjects was put in a separate room. Twenty-five pairs of subjects each contributed eight races: four races under a cooperation condition and four under a competition condition (total of 200 races). Subjects were kept blind to the condition but were aware that some races were cooperative and some competitive. Target horses were randomly assigned to each subject. It was predicted that the cooperation scores would be higher than the competition scores. The result was opposite to that expected: $t_d(24) = -2.214$.

Relatively little experimental work deals directly with the question of whether PK is additive. There are, however, a few reports that might be considered to bear on the issue. Humphrey (1947) and Feather and Rhine (1969) conducted dice experiments in which two subjects were sometimes in a helping condition and sometimes in a hindering condition. McMahan (1945, 1946, 1947) held "PK parties" at which small groups of people tried to influence the falling of disks, although there was no comparison between group and individual performance.

Others investigated situations where one subject's psi influence might be expected to interact with that of others. Schmidt (1985) examined this in the context of retroactive PK. Braud (1985) experimented with blocking psi influence. Some looked at the "checker effect" (e.g., Weiner & Zingrone, 1986) in studies that investigated the psi influences of both subjects and data checkers. I will not attempt to integrate these findings because they are diverse methodologically and based on a variety of underlying theoretical ration-

An earlier version of this paper was presented at the Parapsychological Association 28th Annual Convention, Medford, MA, August 12–16, 1985.

I wish to thank Rick E. Berger, Debra H. Weiner, Rex G. Stanford, and Ephraim I. Schecter for comments on earlier drafts of this paper.

ales. I mention these few references for researchers who may wish to ponder the issues.

In this paper my interest is with the possibility of humans' psi adding together in real time. My hypothesis is that psi influences will add. This idea is not derived from any articulated theoretical perspective but advanced simply as an interesting possibility.

Stanford (1977) suggested that any work on the question of whether PK is additive should be done under conditions where subjects are blind to how many people are trying to influence an event. The present experiment implements his suggestion. It uses simulated horse races in which pairs of subjects attempt to influence the outcome of the races. The races are controlled by an electronic random number generator (RNG) in real time. Each subject is in a different room with an assigned list of target horses and observes the races on a video monitor connected to an Apple computer. Half the time subjects are cooperating (i.e., trying to influence the same horse), and half the time they are competing (trying to influence different horses). Game-like testing situations have been used in a number of psi studies, and Gissurarson (1986) has provided a review.

It should be noted that the terms *cooperation* and *competition* are not used here in the same sense as they have been used in some other parapsychological reports. In this experiment, there is no cooperative or competitive psychological "set" for participants because they are blind to the experimental condition.

Hypotheses

The experimental procedure allows a direct comparison of the psi performance of individuals and pairs. It was predicted that the effects from pairs (cooperation) would be stronger than those from individuals (competition). For purposes of this experiment, "stronger" means higher scoring. To test this, it was decided in advance that two conditions must be met. First, the critical ratio of the difference (CR_d) between the cooperation and competition scores must be significantly positive with alpha set at .05 (a one-tailed test was prespecified). The second criterion was that the dependent t (t_d), using the subject pair as the unit of analysis, must also be significantly positive. Fulfilling the first criterion would assure that the overall cooperation scoring was higher than the competition scoring.

The t-test criterion is included because Stanford and Palmer (1972) have argued that meeting it would help assure that the results are generalizable and that the effect would not be due to high scoring of just a few subjects. Meeting only the t-test criterion, however, would indicate that the value of the t statistic was due, in part, to an empirical variance less than the theoretically expected variance. Such an outcome could pose problems for a direct interpretation. Therefore, the CR criterion is also included.[1]

A secondary hypothesis (of less interest to me, the experimenter) was that the variance of the cooperation scores would be larger than that of the individual competition scores (using the race rather than the subject or pair as unit of analysis). This had been previously found by Feather and Rhine (1969) and Hansen (1982).

Method

Subjects

Twenty-five pairs of people contributed two sets of four races each. The number of subjects was specified in advance. They were drawn from the FRNM participant pool, friends of the experimenter, staff, and students at the FRNM. During recruitment, volunteers were asked to bring along someone with whom they felt comfortable. No one was given a partner; they were required to bring their own. No person was a member of more than one pair, and the experimenter served as a subject as a member of one pair.

Hardware

An Apple II+ computer with 48K memory, two disk drives, one Amdek I color monitor, and two identical black and white monitors were used. The monitors were connected in parallel so that the same information was displayed on each. The source of randomness used to control the movement of the horses was provided by a Bierman-RIPP RNG (for a description see Bierman, 1982). The Apple also contained a Mountain Computer CPS card and an 80-column card (neither of which was used in the experiment).

[1] The CR_d was calculated in the same manner as the t_d except that the theoretical variances (from Monte Carlo simulations) were used rather than the empirical variances.

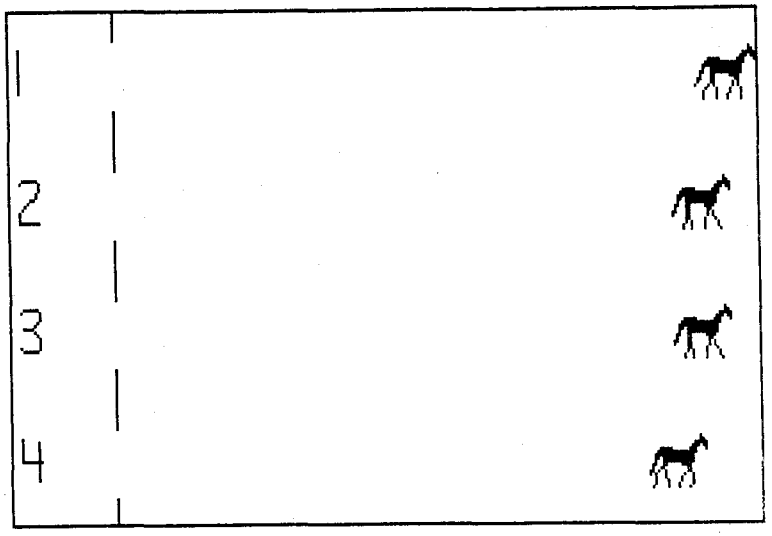

Figure 1. Video display of horse race.

Setup

The experiment was conducted in three rooms on the second floor of the Institute for Parapsychology. One room housed the Apple, disk drives, and color monitor. A black and white monitor, approximately 15 feet directly south of the computer, was housed in a second room. A third room housed the other black and white monitor, which was approximately 30 feet southeast of the computer. The doors to all three rooms were kept shut during the races.

Program Description

The computer program controlling the psi test displayed four horses on the screen (see Figure 1). The horses moved from left to right making one step at a time; the rate was controlled by the RNG. The first horse to cross the finish line (having taken 120 steps) was the winner. After a horse won, the race stopped, data were stored on diskette, and an encouraging message was put on the screen. Each time the program was run, four races were generated.

The horses were drawn and animated using the *Graphics Magician* software package by Penguin Software. The legs and tails of the horses moved with respect to their bodies. Starting and finishing lines were also drawn on the screen. The program was written in

Applesoft BASIC. The RNG was sampled by PEEKing Slot 7. All results were stored on diskette. Bit-level data were stored so the entire race could be recreated step by step.

When the program was initiated, it requested names of the participants and the code numbers of their assigned target lists. After these were input, the experimenter hit the return key, and four races were generated and stored with no further intervention by the experimenter. For each "cycle" of the race, the RNG was sampled, and the most significant bit was checked. If a 1 was obtained, the horse advanced one step; if a 0 was obtained, the horse stayed in the same position. After this was done for all four horses, the program checked to see if any horse had accumulated 120 steps. If any had, the race was finished; if not, another cycle was generated. The program allows more than one winning horse, but this happens only occasionally. The RNG was sampled approximately 14.5 times per second; the average length of a race was about 61 seconds (the length varied slightly from race to race). Between each race there was a pause of approximately 35 seconds.

Assigned Targets

Before each set of four races, subjects were given a list in an envelope indicating which of the four horses they were assigned for each race. The lists were different for each participant. For two of the races in a set, the subjects were assigned the same horse (cooperation) and for the other two races, the subjects were assigned different horses (competition). The order of the cooperation and competition conditions was randomly determined for each set. Subjects were kept blind to whether a particular race was in the cooperation or competition mode.

These assigned targets were all determined before any data were collected. Targets were selected using a noise diode-based RNG in the FRNM DEC PDP 11/45 computer. A pair of lists was generated for each set of races. The first step of the procedure resulted in lists of four digits ranging from 1 to 4; this designated one target horse for each of four races (an open-deck, with replacement method was used). After these lists were generated, the RNG designated, for each list, two of the four digits that would be used for the cooperation trials; the other two would have a different pair of numbers selected for the competition condition for the associated list. Several additional pairs of lists were generated in case of aborts or other problems. The lists that were ultimately used were checked for first-

order dependency (excess or deficit of a digit), and none was found (chi-square values for the cooperation targets and two sets of competition targets were 4, 1.36, and 1.04, each with 3 df). After the lists had been generated, the experimenter copied them onto separate sheets, recorded the code number on the sheets, placed each sheet into an envelope, and marked the code number on the envelope. The code number indicated the sequence and the specific experimental room in which the lists were to be used. The experimenter did not look at any of these lists again until the end of the experiment (the subjects were requested to replace the lists inside the envelope before returning them to the experimenter).

Instructions to Subjects

The instructions varied because of the vast range of sophistication vis-à-vis parapsychology. For some, this was their first visit to the Institute; others had participated in a number of psi experiments; several had spent years doing professional work in the field.

For relatively naïve subjects, prior to the experiment, I spent 15 to 45 minutes discussing parapsychology, psychic functioning in general, and PK in particular. Various types of PK were mentioned and examples were described. I explained the distinction between micro- and macro-PK. It was stated that everyone may have some PK ability and that this might be more widespread than realized.

The exact nature of the experiment was then presented. I explained that they would be trying to influence the outcome of a computerized horse race and that sometimes they would be cheering on the same horse and sometimes different horses but they would not know which races were competitive and which were cooperative. I mentioned that it was unclear whether PK abilities could be added together, but that some earlier research had indicated that they might be additive. I explained that more evidence was needed before a definite statement could be made and that this was the reason for the experiment. It was noted that if PK could be added, it may have practical implications, such as for healing.

To further stimulate enthusiasm for the experiment, it was mentioned that the Institute was looking for successful participants to help in other studies. I noted that we ultimately hoped to form psychic teams to compete with other labs via computer hookup and that we might eventually hold a "psychic Olympics."

After the discussion, participants were taken up to the room containing the Apple computer. A demonstration of the program was given. This consisted of running a full set of races with a sample target sheet listing assigned horses. The demonstration races were controlled by the RNG, and thus the outcomes varied. During this demonstration, I discussed various strategies that might be used such as yelling at the horses, pounding on the table, smearing spittle on the screen in front of the target horse. I explained that I had observed parapsychologists use such strategies in other studies. Several of the participants sang to the horses. More passive wishing strategies were also discussed. My impression is that a more active, striving set was used by most participants, as though they were observing a real horse race, although I did not try to measure this.

After the demonstration, any further questions were answered, and the subjects were given a consent form to sign. The participants were then taken to the other experimental rooms and given an envelope containing the list of assigned targets.

Procedure

After the subjects were situated in the experimental rooms, the experimenter returned to the room with the Apple computer and initiated the program. After inputting the names of the participants and the assigned target-list code numbers, the experimenter covered his video screen and stayed in his room (except when serving as subject), and four races were generated. The experimenter then recorded the date and time in his notebook and made any other brief notes that seemed pertinent.

At the end of the first set of four races (announced by a buzzer), the experimenter retrieved the target lists from the participants and then gave them a second list. The experimenter returned to his room and initiated the second set of races. At the end of this set, the lists were retrieved and the participants thanked for their help. No additional feedback was given to the subjects at this time because the experimenter wished to remain blind to the results until the end of the data collection. The participants were told that at the end of the experiment they could expect to receive a letter containing a written record of their results.

Data collection occurred from May 8 through July 31, 1984. The author served as experimenter for the entire study.

Two sets of races were aborted. For one, a monitor in one room did not function. For the other, a power interruption caused by a thunderstorm crashed the program part way through the set. Before the data were examined, it was decided not to include these sets in

the formal analysis. The participants in both cases contributed an additional set of races.

It should be noted that subjects had no physical access to the computer during the races because they were in other rooms. Only the experimenter initiated the races; subjects had no control over this. The participants only observed the video output and were the first observers of the race outcome in conjunction with the assigned targets. Thus, if the results are interpreted with the subjects as psi sources, this study may be considered as having a PK experimental design rather than a precognition or intuitive-data-sorting design.

Statistical Considerations

The distribution of the race scores is not normal. At least one horse in each race achieves a score of 120. To determine the actual distribution of the scores, a Monte Carlo simulation was conducted with 10^7 races (producing 4×10^7 scores). The simulation was done with several Apple II series computers with several different Bierman-RIPP RNG boards. A BASIC program and an assembly language routine were used. The histogram of the resulting scores is shown in Figure 2. The mean score was found to be 112.176 (± 0.0023, 95% confidence limits) with a standard deviation of 7.377.

For the result of a given competition race, the mean score of the two assigned horses was determined. A second simulation of 10^6 races was conducted to determine this distribution. The mean was found to be the same as the single-horse distribution (within chance fluctuation). The standard deviation was found to be 5.134. This is slightly less than what would have been expected had the scores been completely independent. The scores within any one race are not entirely independent. At least one (and usually only one) horse achieves a score of 120.

Although the underlying distribution is not normal, the *t* and *CR* statistical measures are appropriate because of the large number of races (100 cooperation and 100 competition). This was verified by simulating 50,000 times the entire experiment (all 200 races) using a Bierman-RIPP RNG and an assembly language routine. A distribution of differences between total competition and cooperation scores was determined.

Randomness Tests

The primary check of the RNG was a matched set of control trials conducted at least 2 hours after each experimental session. The experimenter ran the two sets of races without observers. The program was essentially identical to the one used for the experimental condition except that, on initiation, the program requested the date and time be entered. All the same peripheral equipment was in place and turned on as in the experimental session. The experimenter left the room during these sessions.

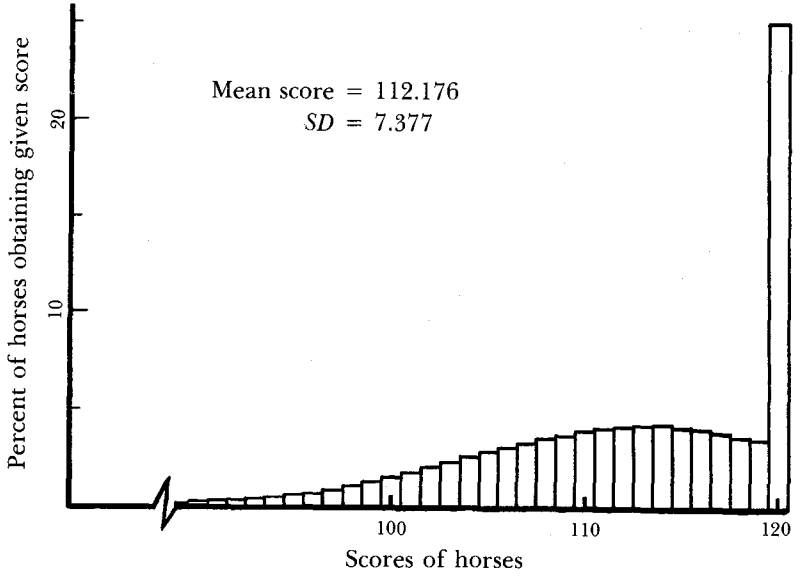

Figure 2. Histogram of scores from Monte Carlo results, with 10^7 races simulated.

The results were analyzed by assigning the same set of target horses that had been used in the corresponding experimental set. This allowed the control data to be analyzed by the automated data analysis program. It was found that the mean of corresponding "cooperation" condition was 112.23 and the mean of the corresponding "competition" condition was 112.19. When the "participant pair" was used as unit of analysis, the paired $t(24) = 0.04$. The standard deviation of the "cooperation" race scores was 7.57 and that of the individual "competition" race scores was 7.24; $F(99, 199) = 1.09$.

In addition, after the control trials were run, a slightly modified and compiled version of the RNG test for randomness given by Bierman (1982) was run. This collected 10,000 bytes and performed a chi-square test on these bytes (255 *df*). Twenty-six of these tests were run. The largest chi-square value obtained was 300.36; the smallest

was 219.30. Thus, none were statistically significant at the .05 level, two-tailed. A Kolmogorov-Smirnov test was performed on the chi-square values by using the random analysis disk supplied with the Psychophysical Research Laboratories' (1985) PsiLab // package (K+ = .775, K− = .262). The results are well within what is expected by chance.

It should again be mentioned that the assigned target horses, for both experimental and control conditions, were randomly selected in advance by a different random procedure, as previously described. Also, the analysis uses differences of scores. These methods give yet further assurance against an artifact that is due to bias of the RNG.

Results

A computer program was written to read the original data stored on one disk and a file with the assigned targets on a separate disk and then compute the CR_d and t_d of the hypothesis. A supplemental check was made by hand, which was done by printing out all the scores of the races and then tabulating the scores of the assigned horses. This verified the operation of the analysis program.

The overall results are shown in Table 1. As can be seen, the competition scores were higher than the cooperation scores: $t_d(24) = -2.214$, $CR_d = -2.058$. The Monte Carlo simulation of 50,000 experiments indicated a one-tailed p value of .9800 ± .0012 (95% confidence interval); this corresponds to $CR_d = -2.054$. The hypothesis was not confirmed; in fact, the results argue against it. If a two-tailed test had been designated in advance, the result would have been significant. Given the hypothesis, however, a test with unspecified direction would seem to make little sense.

The standard deviation of the cooperation races was 7.11, and the standard deviation of the individual competition race scores was 7.25, $F(199, 99) = 1.04$, which was also opposite the predicted direction. These variances were calculated using the empirical means rather than the theoretical means (the theoretical mean was used in the calculations of Feather and Rhine, 1969). One needs to be cautious in using an F test under these circumstances because there is a small dependence between competition scores within a race and the underlying distribution is not normal. However, it is clear that the result is close to chance expectation.

Table 1
Mean Scores, with Each Value Based on Four Races

Pair	Sex	Cooperation	Competition[a] Person A	Person B
1	F-M	109.25	116.50	108.75
2	F-M	110.50	108.50	113.25
3	M-F	107.25	117.25	112.50
4	F-F	104.75	103.00	119.50
5	F-F	117.50	114.75	109.50
6	M-M	111.25	112.00	108.75
7	M-M	114.00	118.00	108.25
8	F-M	108.75	113.00	111.75
9	F-F	113.25	114.75	107.50
10	F-M	106.00	116.50	113.50
11	M-F	114.00	118.75	112.75
12	F-M	113.25	113.50	114.75
13	F-F	107.25	114.50	115.75
14	M-F	118.50	106.50	115.25
15	F-M	111.25	118.00	107.25
16	M-F	112.25	107.00	113.75
17	M-F	109.75	113.25	109.50
18	F-M	109.75	108.25	118.75
19	F-F	108.25	118.75	108.75
20	F-F	107.25	110.50	106.75
21	M-F	116.75	115.00	115.50
22	M-F	112.75	114.25	119.25
23	F-F	113.00	110.00	111.75
24	M-F	112.25	115.50	114.00
25	M-F	107.75	114.00	116.50
Overall mean		111.06	112.91	

[a]The order of the competition scores corresponds to the sex list (e.g., for Pair 1 scores, Person A is a female and Person B a male).

Post Hoc Analyses

The following results are reported so that future quantitative reviews might be facilitated.

The predictions for this experiment only involved differences between conditions. Differences from theoretical chance distributions are reported in this section. The cooperation scores were below MCE, $t(99) = -1.56$; the individual competition scores were above MCE,

$t(199) = 1.44$. Both these scores were calculated using the race score as unit of analysis. This report contains sufficient information for the interested reader to calculate *t*'s using the subject as unit of analysis and for an overall *CR*.

For the horses that were *not* targets, in the cooperation condition, the mean was 112.34, $SD = 7.25$, $N = 300$. For the competition condition, the mean was 111.66, $SD = 7.27$, $N = 200$.

Rao, Kanthamani, and Krishna (1979) had found that in ESP testing under competition conditions, one person would do well and the other poorly, that is, there would be a large difference between their scores. To test this idea, the absolute differences of the two competition scores for each race was determined. The mean of the differences was found to be 8.8. A simulation of 10^6 races was run; the mean was found to be 8.51 with $SD = 6.30$ ($CR = .46$). It should be noted that the methodology of the Rao et al. study was drastically different than the present experiment.

Discussion

The hypotheses were not confirmed. Too many parapsychology reports offer speculative ideas as to why hypotheses are not supported. I do not wish to contribute to that trend. I do not wish to interpret these results or to offer any rationale for the outcome. Furthermore, I will make no argument that psi did indeed "occur" or that it did not "occur." It remains to be seen whether these results will fit into any patterns uncovered in other research.

References

Bierman, D. J. (1982). *PSI-test kit random number generator for the Apple.* Amsterdam: Research Institute for Psi Phenomena and Physics.

Braud, W. (1985). Blocking/shielding psychic functioning through psychological and psychic techniques: A report of three preliminary studies. In R. A. White & J. Solfvin (Eds.), *Research in parapsychology 1984* (pp. 42–44). Metuchen, NJ: Scarecrow. (Abstract)

Feather, S. R., & Rhine, L. E. (1969). PK experiments with same and different targets. *Journal of Parapsychology*, **33,** 213–227.

Gissurarson, L. R. (1986). RNG-PK microcomputer "games" overviewed: An experiment with the videogame "Psi Invaders." *European Journal of Parapsychology*, **6,** 199–215.

Hansen, G. P. (1982). A cooperation/competition PK experiment with a computer simulated horse race. *Journal of Parapsychology*, **46,** 57–58. (Abstract)

Humphrey, B. M. (1947). Help-hinder comparison in PK tests. *Journal of Parapsychology*, **11,** 4–13.

McMahan, E. A. (1945). PK experiments with two-sided objects. *Journal of Parapsychology*, **9,** 249–263.

McMahan, E. A. (1946). A PK experiment with discs. *Journal of Parapsychology*, **10,** 169–180.

McMahan, E. A. (1947). A PK experiment under light and dark conditions. *Journal of Parapsychology*, **11,** 46–54.

Psychophysical Research Laboratories. (1985). *PsiLab II*. Princeton, NJ: Psychophysical Research Laboratories.

Rao, K. R., Kanthamani, H., & Krishna, S. H. (1979). Psi and volitional activity: An experimental study of psi in a competitive setting. *Journal of Parapsychology*, **43,** 101–112.

Schmidt, H. (1985). Addition effect for PK on pre-recorded targets. *Journal of Parapsychology*, **49,** 229–244.

Stanford, R. G. (1977). Experimental psychokinesis: A review from diverse perspectives. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp. 324–381). New York: Van Nostrand Reinhold.

Stanford, R. G., & Palmer, J. (1972). Some statistical considerations concerning process-oriented research in parapsychology. *Journal of the American Society for Psychical Research*, **66,** 166–179.

Weiner, D. H., & Zingrone, N. L. (1986). The checker effect revisited. *Journal of Parapsychology*, **50,** 85–121.

Princeton Arms North 1, Apt. 59
Cranbury, NJ 08512