Amplifying Precognition
Four Experiments with Roulette*

Dr. Wilfried Kugel
Freie Universität Berlin, Fachbereich Physik, Institut für Experimentalphysik

Kogel

The Parapsychological Association
42nd Annual Convention

August 4-8, 1999
Stanford University
Palo Alto, California

Program Committee
Kathy S. Dalton, Chair

Referees

Cheryl Alexander
Carlos Alvarado
Dick Bierman
Rick Berger
Richard Broughton
Norman Don
Stanley Krippner
Bruce McDonough
Robert Morris
Roger Nelson
John Palmer
Dean Radin
Ephraim Schechter
Simon Sherwood
Matthew Smith
Jerry Solfvin
James Spottiswoode
Fiona Steinkamp
Russell Targ
Charles Tart
Marcello Truzzi
Jessica Utts
Caroline Watt
Nancy Zingrone

Arrangements Committee

Adrienne Smucker
Marilyn Schlitz
Dean Radin
Edwin May

Amplifying Precognition Four Experiments with Roulette*

Dr. Wilfried Kugel
Freie Universität Berlin, Fachbereich Physik, Institut für Experimentalphysik

Abstract

Starting point of the investigation is the well known fact, that subjects asked to produce random decisions are not able to do so, but will show characteristic patterns in their choices. This behavior is influencing the results of statistical ESP-experiments (here: precognition). Using an information theoretical model, a computer program was prepared to recognize such patterns and match it with the outcome of the subject's precognition trials. The program was able to make its own bets on certain trials when the appropriate statistical criteria were met. The programs bets did not have to be the same as the subjects bets.

Four precognition studies were completed between 1980 and 1999. To motivate the subjects, the experiments were performed as roulette games. Comparing the subjects' hit rate with the program's, the program performed considerably better, "amplifying" precognition.

The overall score of the amplifying program resulted in a highly significant z-value of +2.87 ($p = 0.002$), where the subjects scored below chance.

Introduction

It is an old dream to find a method of predicting the outcome in games of chance.** Several attempts have been made to amplify subjects' possible precognition abilities in game situations: In the 1950s Milan Ryzl (1971) performed experiments to predict the winning numbers in a lottery, and developed a (complicated) statistical method to amplify the subject's performance. In the 1960s, Robert Brier (1970) together with the publisher of the journal *Rouge et Noir* performed some experiments at the Casino of a US hotel (Brier & Tyminski, 1970), also trying to amplify the performance of subjects with the help of a statistical method. Both claim to have won money with the help of their methods.

Under the non-ESP hypothesis, the outcome of a game of chance is completely random, no matter what strategy the subject applies. Under the assumption that precognition might be effective, future

* A previous version of this paper, covering the first two studies A and B, has been published in the EJP (see Kugel 1990-1991). Study D was funded by the *Institut für Grenzgebiete der Psychologie und Psychohygiene e.V.* (Freiburg i.Br.).

** French aristocrats gambling passions and the arising questions led Blaise Pascal (1623-1662) and Pierre Fermat (1601-1665) to the foundations of the theory of probability.

events must correspond to the subject's behavior in the presence (and past). Thus, in the case of roulette, random events should be detectable in the subject's response behavior. Furthermore, if the subject acts in a way that is completely strategy-dominated, it can be assumed *a priori* that no precognition is involved.

Knowing that a betting game seems to be a high motivating test situation for subjects, and that it might be promising to filter the information output at the subject's side (bets) with the help of statistical methods, led the author to develop a computer program that can perform real-time calculations during a roulette game and so may able to amplify precognitive information in a betting situation.

The mathematical formalism, developed with the help of methods of the information theory, was first introduced by the author at the 1978 PA convention in St. Louis, Missouri (Kugel, 1979).

The results of two series' of experiments (A: 1980, B: 1981-1984) were reported in the 1990-1991 volume of the EJP (Kugel, 1992).

This current paper, for the first time, reports the results of two further confirmation studies (C: 1992, D: 1997-1999).

Response Patterns

Hans Reichenbach (1935) stated that mathematically naive subjects would not be able to produce a random series of alternative symbols taken from a set. This hypothesis was confirmed in several psychological investigations (see: Tune, 1964).

In test situations, subjects when forced to produce random decisions, show preferences and aversions. Furthermore, the subjects avoid repetitions of the same choice (Mittenecker, 1958). Reinforcing the response (feedback) leads the subjects to behave in the opposite way; most of them prefer to repeat the same choice.

Thus, every chain of the subject's responses shows a well-proportioned structure, characteristic for the individual like a fingerprint and determined by a strategy. If one takes the subject's decisions as output of an information receiver, it is clear that transmitted information cannot manifest itself when the subject's behavior is completely guided by a non-variable, pre-formed pattern; to speak in terms of radio technique, the receiver has no "oscillator circuit".

As is known from quantitative and qualitative ESP-Experiments (Tenhaeff, 1976), correct readings as well as high hit scores seem to depend on individual habits of the subjects. Such habits might be imagined as reasoning in the personality structure of the subjects. Habits manifesting themselves in the form of response patterns in statistical ESP experiments were investigated by several experimenters (Goodfellow, 1938; Martin & Stribic, 1938; Stanford, 1967; Morris, 1971; Mischo, 1972; Kugel, 1977).

Measuring Response Patterns

Every hypothesis about response patterns has to be transformed into a hypothesis about relative frequencies of such patterns in the response series. Information theory is - especially in real time experiments - a powerful tool for calculation of values which represent the strength of strategy application.

A functional U was defined, which is always non-negative and increases with the divergence between two sets of probabilities (or relative frequencies).This *U*-value is independent of the number of trials; it rises and falls with the actual divergence between two sets and can be used to measure patterns in the subject's response series with respect to a given reference. It can easily be calculated in real time (for details see: Kugel, 1979).

Three applications of the U-measure on the two alternatives case (binary choice) were used:

(1) *Frequencies (transition probabilities of the Order 0).* In a decision situation, every subject tends to prefer or avoid one of the alternatives. In the extreme case of a strong response strategy, only one alternative is chosen.The measure U_a detects the favoring/avoiding of alternatives.

(2) *Frequencies of "equal and unequal pairs" (transition probabilities of the Order 1).* Investigating the alternation of choice, one finds that the subjects usually alternate much more than expected by chance. They avoid repetitions of the same choice.The measure U_b detects whether more pairs of unequal responses (repetition avoidances) or pairs of equal responses (repetitions) appear.

(3) *Feedback response*. Positive feedback can be seen as a reward. Thus, the rewarded choice creates in the subject a tendency to repeat that choice. But feedback is an external stimulus which contradicts the internal stimulus for "repetition avoidance". Feedback thus causes two different effects. On the one hand, feedback-triggered response is a strategy, and as a strategy it weakens the potential manifestation of psi-mediated information on the subject. On the other hand, feedback response weakens the quite dominant strategy of "repetition avoidance". This double effect may very well explain the controversy of whether or not feedback increases PSI performance. The measure U_c detects whether or not the subjects respond to feedback.

Implementation into a Program

In a betting situation, for every single trial *j*, 3 values can be measured: subject's prediction (1 or 2 with regard to the colors *noir* or *rouge*), subject's wager on this prediction and the colour of the random number (1, 2 or 0) generated after the subject's bet. After a certain number of trials, enough values have accumulated to calculate the 3 *U*-values mentioned above, using the data from all previous trials *(1...j)* with the help of a computer program. For the actual trial *j* it can be determined, whether or not the *U*-values are rising or falling with respect to the preceding trial *(j-1)*:

$$U_a(j) > U_a(j-1) \text{ or } U_a(j) \leq U_a(j-1)$$
$$U_b(j) > U_b(j-1) \text{ or } U_b(j) \leq U_b(j-1)$$
$$U_c(j) > U_c(j-1) \text{ or } U_c(j) \leq U_c(j-1)$$

This leads to $2^3 = 8$ different possible strategies. For each of the 8 strategies, the number of corresponding hits (positive sign) and misses (negative sign), weighted with the related (square* of)

* The wager can be seen as a variable indicating how sure the subject was about the single prediction. The square of the

subject's wager were stored in the program's session memory.

For the forthcoming trial *(j+1)*, we can now calculate both alternatives of changes in the *U*-values with respect to both predictions, one could make.[**] Thus, one can decide which of the possible two predictions would represent the better strategy, with regard to the effectiveness of all strategies applied by the subject during the previous trials. The program's prediction for the forthcoming trial does *not* consider the subject's prediction for the forthcoming trial, but acts completely independently, only on the basis of the data of the previous trials.

In the program versions from series B on, dynamic software filters were applied to suppress random fluctuation. The filters suppressed predictions, when not all values were defined (for instance during the build-up of statistics in the beginning of the session) or when some values were (nearly) equal to each other.

Table 1: Flow Diagram of an Experimental Session

Chain of Events in a Session
Intialization
*statistics for alternative 1 (U-values)
statistics for alternative 2 (U-values)
selection on the basis of the more efficient strategy
filter
strategy values stability control, termination at instability
generation of program's bet, if possible
subject's game (game action)
imput of subject's bet and wager
control for nput errors
update of program variables
program's bet, if possible (game action)
throw (game action)
input of throw
control for input errors
update of program variables
partial initialization – continue at *

wager was used to amplify the differences between wagers from trial to trial.

[**] In program version 1 (study A) for every trial a total of the three U-values was calculated. The program then always predicted the colour for which $U(j+1) \leq U(j)$. For the first trials, when not enough data for calculation were available, the program simply took the same choice as the subject. Since program version 2 (study B), the three U-values were handled separately.

TERMINATION CONDITIONS

The only termination condition for study A was the completion of 20 sessions of 50 trials each. Termination conditions for the session in study B consisted of: program errors, input errors by the experimenter, the session time being longer than 1 hour, the real game performed faster than the calculation time of the computer, unstable strategy values, an uncomfortable situation at the casino, or request from the subject to stop. Of course, this means "optional stopping". But at the casino, we were only in a "study" and no limits for either the number of trials or for the program's predictions were set.

The only termination condition for study C was the completion of 500 program's prognoses, for study D the completion of 2.000 program's prognoses.

Additionally, in study D, only the first 28 trials of single sessions contributed to the result. This condition was set before the beginning of the series to avoid decline effects. Furthermore, in study D, the time for a single session should not exceed 45 minutes, and the subjects could terminate the session whenever they wanted.

THE EXPERIMENTS

Study A (Technical University Berlin, 1980)

This study performed at the Technical University Berlin PSI-lab. The results were first published in the annual research report of the project (Kugel, 1980).

In September 1980, the author introduced a computer program (version 1, in FORTRAN) to predict even chances in a roulette game, using the *U*-values as predictor variables. A toy roulette device* was used. The program was run via telephone modem on a CD CYBER 172 of the Technical University Berlin computer centre.

Twenty subjects were tested, only one session per subject was allowed, and each session consisted of 50 trials. The subject could bet any (hypothetical) wager on *rouge* or *noir*. During the first trials of a session, the program had to build up its statistics: it simply made the same predictions as the subject.** After the variables were defined, the program started to calculate its own prediction (not taking the subject's actual bet into account) and made its bet after the subject had done so. The program always did bet the same wager as the subject, no matter whether its prediction differed from the subject's prediction or not. Then the experimenter threw the ball. When a 0 appeared, the ball had

* The roulette wheel had no manufacturer listed other than "Made in Italy"; the wheel was approximately 12" (30 cm) in diameter.

** This method was used because the same number of predictions was wanted for the subjects and the program. (There was no bet of the subject without a corresponding bet of the program.) Of course those trials where the program was just placing the same bet as the subject should not be included in the analysis. But for this experiment, we had to do so because it was fixed in the methodological set up prior to the experiments (to make the statistics easier). The hypothesis only stated that the program would have more hits in the majority of sessions. Under this hypothesis it does not matter if subject and program make the same bet for a while because the difference of program's and subject's hits in these cases equals 0.

to be thrown once more.

There were two hypotheses:
H1: In the majority of sessions the program's hit score will be higher than the subject's.
H2: In the majority of sessions the program's winnings*** will be higher than the subject's.
In 13 sessions, the program performed better than the subjects, and in one session, both obtained an equal number of hits. Only in 6 cases did the subjects perform better than the program. (H1: $p=0.084$, one-tailed)

In 14 out of 20 sessions the (hypothetical) winnings of the program were higher than those of the subjects (H2: $p=0.058$, one-tailed). Though the hypotheses could not be accepted, the results looked promising.

*Study B (Berlin Casino, 1981-1984)**

In 1981 programmable pocket computers became available in Germany. A shorter version of the original program was developed (version 2, in BASIC) for the Sharp PC 1211 (with 1.4K memory). It was very difficult to fit the program into the very small memory, so only the most important calculations could be made. The actual decision about the prediction (on the basis of the 3 U-values) had to be done by the experimenter during the session. The computer's calculation time for one trial was about 90 seconds. In some sessions, because of the long calculation time, the game at the casino table ran faster than the calculation time and the session had to be terminated. Several program versions (2-4) for the Sharp PC 1211 were developed. Since 1982 a Sharp PC 1500 (with 10K memory) was used. The calculation time could be lowered to 7 seconds. For this computer, program version 5 was developed.

The subjects got their total wagers (300 DM, later 600 DM**) from the experimenter. They had no financial risk, because they could not lose any of their own money. Additionally, they got half of the session's net win from the experimenter. The subjects could bet any sum on even chances. The experimenter fed all the data into the computer. When the program made predictions, the experimenter made this bet *after the subject had placed the bet*, thereby reducing the influence of the program on the subject.

Working with 5 selected subjects at the Berlin Casino, 39 sessions were performed from 1981 to 1984. In 5 sessions the program could not make any predictions. In 6 of the remaining 34 sessions, program and subject had the same number of hits. From the now- remaining 28 sessions, in 15 cases the program performed better than the subjects, which is near chance. But the total number of the program's hits was, with 17 above chance expectation, significantly raised.

*** Winnings = monetary prizes (toy money).

* This study was sponsored by a private person.

** The amount of money the subjects could play with had to be raised because with only 300 DM, in some sessions, not enough games could be performed to use all predictions of the program: The session would be terminated before because the subject went bankrupt.

Study C (Berlin, 1992)

Study C was performed in 1992 at the home of the author. A computer program (Version 6) was developed for the PC, simulating the roulette game at the screen. A true random number generator was constructed, using a radioactive uranium source.

54 sessions with about 50 subjects werde performed. The subjects could bet any (virtual) wager on *rouge* or *noir*. The program made its bet only after the subject had done so.

There was only one hypothesis, that the hit score of the 500 prognoses of the program would be significantly higher than chance expectation. This hypothesis could not be confirmed, but the deviation was lying in the expected direction.

Because of a lack of funding, the hardware used in series D was of very poor quality. Furthermore, the program consisted of a number of modules programmed in QBASIC, TBASIC, C and FORTRAN77. There were several technical problems during the experiments. But fortunately, the random number generator did work well.

Study D (Freie Universität Berlin, 1997-1999)

Study D was designed to confirm the results of the previous 3 studies with much better equipment and under much better conditions. This study was performed at the physics department of the Freie Universität Berlin. Program version 7 was developed in Borland C++ for Windows 95.

Now, 2,000 program prognoses, more than all in all three previous studies together, were generated. This took 389 sessions and a total of 10,641 single trials.

The main hypothesis was, that the hit score of the 2,000 prognoses of the program would be significantly higher than chance expectation. This main hypothesis could be confirmed.

For this study, a second computer system was build up:
The hardware and software was completely equal to the main system. This second system received via a radio mouse and an infrared keyboard completely the same input from the subjects as the first system. But it was equipped with an own Geiger-Müller counter which was turned 180 degrees with reference to the Geiger-Müller counter of the first system. So, the second system received different energy quanta from the source, resulting in a different random number string. There was no feedback of the second system to the subjects, who saw only the screen of the first system. Thus, the second system represents a situation with false feedback (since the subjects received feedback from the first system). No significant deviations from chance expectation could be observed at the second system. Especially, the prognosis program running at the second system could not reach an above-chance result.

There was no sign for a correlation between the two random series' of the two systems, nor did the target series' show any deviation from chance. So, there is some evidence that in series D, there was no PK influence.

Results

Random Numbers

In series A, a cheap toy roulette device was used. Surprisingly, the randomness of this device appeared to be very good. Series B was performed in a Berlin casino. So, real roulette devices were used, which were controlled officially.

In series' C and D, a radioactive source (Uranium 238, contained in natural UO_2) was used as random source.

The time interval between two events (peaks) of the decay process varies randomly. 38 successive intervals between peaks were measured by a PC. The first interval was always deleted for security reasons (possible "start-up oscillation"). From the remaining 37 intervals, the longest one was determined. The ordinal number of this longest interval (between 1 and 37) minus 1 determined the actual random number (between 0 and 36). In series C and D respectively, different hardware (Geiger-Müller couters) and different software was used. For random analysis, random numbers were reduced to 1, 2 and 0.

For all 4 series', target frequencies and target transition frequencies from lag 1 up to lag 4 were compared to chance expectation. In all 4 series', no significant deviations from chance expectation could be observed. So, all 4 random number generators were working perfectly.

Precognition

All data produced by the prognosis module are given as generated by the program version actually running during the experiments. Although there were errors in program version 6, which were corrected in 1997, the corrected prognosis values are not presented here.

The following program versions of the prognosis module were used: 1 in series A; 2 through 5 in series B; 6 in series C; 7 in series D.

The *a priori* probabilities for a hit were: 1/2 (without Zero) for series A, 18/37 (with Zero) for series' B, C, D.

In all 4 studies, subjects scored collectively slightly below chance, as shown in Table 1. (There was no hypothesis with respect to these scores.)

Table 1: Subjects Results

Series	Trials	Hits	Expectation	Scoring Rate	z Scores	Probability (1 Tailed)
A	900	446	450	49.6%	-0.27	ns
B	945	456	459.7	48.3%	-0.24	ns
C	2,489	1,205	1,210.9	48.4%	-0.24	ns
D	10,641	5,158	5,176.7	48.5%	-0.36	ns
Total	14,975	7,265	7,297.3	48.5%	-0.56	ns

In all 4 studies, the prognosis module scored above chance, as shown in Table 2. (These scores were the main hypotheses for series' C and D.)

Table 2: Prognosis Module Results

Series	Trials	Hits	Expectation	Scoring Rate	z Scores	Probability (1 Tailed)
A	900	459	450	51.0%	+0.6	ns
B	313	169	152.3	54.0%	+1.89	0.029
C	500	255	243.2	51.0%	+1.06	ns
D	2,000	1,022	973	51.1%	+2.19	0.014
Total	3,713	1,905	1,818.5	51.3%	+2.87	0.002

In all 4 studies, synchronous with the program's prognoses, the subjects scored collectively below chance, as shown in Table 3. (There was no hypothesis with respect to these scores.)

Table 3: Overall Precognition Results for Subjects, Synchronous with the Program's Prognoses

Series	Trials	Hits	Expectation	Scoring Rate	z Scores	Probability (1 Tailed)
A	900	446	450	49.6%	-0.27	ns
B	313	152	152.3	48.6%	-0.03	ns
C	500	240	243.2	48.0%	-0.29	ns
D	2,000	942	973	47.1%	-1.39	ns
Total	3,713	1,780	1,818.5	47.9%	-0.99	ns

For all 4 series' together, the overall difference between program's hit scores and the corresponding hit scores of the subjects lies significantly above chance expectation ($z_D = 2.73$, $p = 0.003$, one-tailed).

Discussion

This report covers all roulette experiments ever carried out by the author. All original data and programs have been stored and are available.

The experiments at the Technical University Berlin had to be terminated because the research project was closed in 1981. The casino sessions were terminated because the subjects became increasingly bored by the uncomfortable situation at the casino, which could be regarded as a form of "optional stopping". But the results were promising enough to perform more research.

Two further studies could confirm the results. In all four studies performed so far, subjects scored collectively slightly below chance, whereas the prognosis program scored above chance, in the last 3 studies even significantly above.

This results indicate that a reproducable effect has been found. The algorithm implemented into the program was able to detect an influence of the target information on the subject's behaviour structures. The program was able to use these detected structures to generate its own predictions which appeared to be much more effective than the subject's own predictions.

Under the assumption that the results reported are not caused by chance, which is very improbable, one could interpret the phenomenon observed as strong indication for an information transfer from the future into the past. Even if this information is not consciously available to the subjects, it is detectable by a computer program.

Dr. Wilfried Kugel
Freie Universität Berlin, Fachbereich Physik, Institut für Experimentalphysik
Gast-Forschungsprojekt "Prognostizierbarkeit stochastischer Ereignisse (Präkognition/Psychokinese)"
Arnimallee 14, D-14195 Berlin

Richard-Wagner-Str. 7
D-10585 Berlin
Tel +49-(0)30-3414600
Fax +49-(0)30-3414963
email: wkmail3000@yahoo.de