RESEARCH

Are There Stable Mean Values, and Relationships between Them, in Statistical Parapsychology?

Wolfgang Helfrich

Fachbereich Physik, Freie Universität Berlin, Germany
helfrich@physik.fu-berlin.de

Submitted 12/03/09; Accepted 04/27/10

Abstract—Mean values of the z-scores of statistical psi effects in psychokinesis (PK) and extrasensory perception (ESP) are compiled. Under the influence of psi, the z-score distribution of experiments with binary random number generators (RNGs) at large numbers N of bits is known to be shifted and widened, but to remain indifferent to N. The mean z-score of binary one-trial experiments with dream psi is noted to be not much smaller than the mean z-score of these (presumably) mostly isolated many-trial experiments with RNGs. This could suggest that with two equivalent choices PK and ESP are equally effective, or very nearly so, and the mean z-scores are almost equal at all N down to N = 1. The widening is found to be attributable to a Gaussian distribution of the magnitude of the PK effect that causes the shift, provided the z-score distribution remains Gaussian. Formulas are proposed to compute from the values of shift and widening as observed with RNGs those of psi effects with more than two equivalent choices, such as falling dice. They are only partially confirmed by the (still rather scanty) datasets of such systems. Finally, psi-induced switching and mind-neuron interaction are revisited.

Keywords: statistical parapsychology—mean values—relationships between mean values—psi effects—psychokinesis (PK)—extrasensory perception (ESP)—random number generators (RNGs)

Introduction

Parapsychological effects are notoriously unreliable. In statistical studies of psychokinesis (PK), extrasensory perception (ESP), and mixed psi effects the influence on a single random event is necessarily insignificant, being in the range of statistical scatter. The statistical proof of a psi effect requires large numbers of individual trials, such as bits, falling dice, or guesses. In some cases, an overall chance probability of a psi effect is calculated directly from the data of all existing trials. In other cases, it is preferable or even necessary to begin with analyzing sequences of trials measured on various occasions. A whole sequence is often referred to as an experiment. The number of trials in a sequence may vary from twenty up into the millions. In a meta-analysis, the chance probabilities of the sequences are suitably combined to obtain another overall chance probability. The values of the total chance probability may differ, depending on the method of calculation, but they tend to become smaller and smaller as the number of individual trials or sequences of trials increases. Accordingly, the psi effect becomes more and more credible. Baffling failures and irregularities have occurred in such studies and seem to be typical of psi, but they do not invalidate the overall statistical proofs of existence.

While there are several psi effects whose reality appears to be beyond reasonable doubt, it is not so clear whether some of their properties such as the hit rate in one-trial experiments or the z-score (to be defined immediately) of many-trial experiments have reliable averages. Also, it is not known whether these averages, like quantities in statistical thermodynamics, are related by laws or at least obey rules. Mean values of some psi effects have already been reported, often with fairly small errors. Although their stability in future studies is uncertain, it seems worthwhile to compile them and to look for mathematical relationships between them. This is the purpose of the present article. It is a delicate endeavor because it would probably be a mistake to regard psi effects as a new branch of physics. Psi seems to depend largely on psychological factors, among them talent and mood of the test person or subject, and many other hardly tangible conditions. A brief description of statistical evaluation techniques and an overview of established, statistically evaluated psi effects serve as background and precede the search for relationships.

Outline of Statistical Methods

In the following survey, the statistical studies of psi effects are arranged in groups according to the type of data analysis. In all cases the original aim was to determine either the excess hit rate above its mean chance expectation (MCE) value or the so-called effect size e, i.e. the mean value of the z-score per trial or single random event. These two quantities, which are closely related to each other, are often thought to be constants characterizing a given psi effect. The z-score may be defined as the ratio of the deviation of the number of hits from its MCE value to the root-mean-square (r.m.s.) deviation of the number of hits from MCE. Other designations for the r.m.s. deviation are standard deviation and square root of the variance. With a fixed actual hit rate h, the z-score takes the form

$$z = (h - p)N/[p(1 - p)N]^{1/2} \qquad (1)$$

Here N is the number of trials executed, either in all studies of a psi effect or in the experiment considered, depending on the type of evaluation, h is the actual hit rate, and p is the MCE hit rate, i.e. the probability of a hit in the absence of psi. The effect size e or z-score per trial is defined by the relationship

$$e = z/\sqrt{N}, \tag{2}$$

In combination, Equations (1) and (2) yield for the effect size as a function of h and p

$$e = (h - p)/[p(1 - p)]^{1/2}. \tag{3}$$

On the assumption that they are basically constant, the quantities $h - p$ and, thus, e can be computed directly for any large N from the experimental data. The computed values are expected to have their standard errors. The standard error presupposes MCE and is identical with the root-mean-square (r.m.s.) deviation divided by $\sqrt{N}$. The definition of z is such that its standard error is unity. Although the r.m.s. deviation of z in many-trial experiments has been found to increase under the influence of psi (see below), it is common practice to equate the standard error of the effect size with its MCE value which because of Equation (2) is $1/\sqrt{N}$. Because of Equation (3), the standard error of $h - p$ or h then becomes $[p(1 - p)]^{1/2}/\sqrt{N}$. (The formula for the empirical standard deviation, which refers to the empirical mean value, substitutes $1/\sqrt{(N - 1)}$ for $1/\sqrt{N}$. The difference can be ignored at large enough N.) If the effect size differs among many-trial experiments, an overall mean value is obtained by weighted averaging. Evidently, if the effect size is much smaller than 1, a large number of trials is required to make the standard error smaller than the value to be determined.

If it turns out that $h - p$ and e are not constant but depend on N, one may have to give up the hypothesis of constant effect size. In fact, the mean z-score rather than the effect size appears to be independent of N in a certain class of many-trial PK experiments with binary random number generators. A meta-analysis of a multitude of experimental results has been performed in this case. It makes use of the psi-induced shift and widening of the Gaussian z-score distribution that have been observed in these experiments. The mean z-score of an experiment as represented by the shift can be used to calculate an overall z-score. Details of the meta-analysis, especially possible implications for effect size and hit rate, will be discussed below.

The chance probability of a particular psi effect is, as a rule, derived from its overall z-score. The chance probability P(z) of finding a z-score at or above the argument z is a well-known and tabulated function. In general, a z-score is regarded as significant when $z \geq 1.96$ with $P(1.96) = 2.5\%$. Sometimes, the limit of significance is set at $z = 1.645$ with $P(1.645) = 5\%$.

Experimental Background

Psi Effects Studied in One-Trial Experiments

Let me refer to a book by Radin (2006) for graphic examples of cumulative averaging of the hit rate of some psi effects. A cumulative (or up-to-date) average hit rate covers all studies up to the time when it is calculated. (The attribute "average" will occasionally be omitted, because the hit rate is an average by definition.) In his Figure 6-1, Radin plots the cumulative mean hit rate of dream psi experiments over the year of averaging. He takes most of the data from a review by Sherwood and Roe (2003) dealing with 47 studies. In the dream psi tests, which began in 1966, the subject intends to dream of the contents of a picture selected at random from a given pool of ten. In the early studies, the picture was "sent" by another person during dream (or REM) sleep phases of the subject who was woken up after dreaming to record their impressions. Later on, in a simplified mode, the pictures were displayed in an empty room on a computer monitor while the subject slept, and the subject recorded their dreams from memory in the morning. The next day the pictures in the pool were ranked according to the estimated probability of being the correct one. A hit was registered when the sent picture was in the first half of the ranking. There seems to be no significant difference in the results between the two modes. The final cumulative mean, or overall, hit rate of all 1,270 trials was 59.1%, where 50% is expected by chance. The fluctuations of the early values of the cumulative average hit rate are large (varying between 52% and 65%), which is a consequence of the small numbers of trials involved. When the initial fluctuations had ceased, the cumulative average hit rate fluctuated within the range of (59.1 ± 1)%, and there was no apparent drift.

In his Figure 6-6, Radin (2006) gives a similar graph for ganzfeld psi tests. Here a waking subject in the ganzfeld state is to "receive" a picture or the essence of a video clip. The object is randomly selected from a pool of four and "sent" by another person or only displayed on a computer monitor. After 88 ganzfeld studies with altogether 3,145 trials, performed from 1974 through 2004, the overall hit rate was 32%, while 25% was expected by chance. In addition to the strong initial fluctuations, the cumulative mean hit rate displays, except for the last ten years, a small downward drift starting from 35%. The time taken by a ganzfeld trial seems to have been one or two hours.

Finally, based on studies by Sheldrake (2003), Radin (2004), and others, Radin (2006) in his Figure 6-10 plots cumulative mean hit rates of experiments on the sense of being stared at. Here the subject is to guess after a signal whether or not he or she is being stared at by another person. In 60 different studies, with altogether 33,357 trials, the overall hit rate was 54.5% where chance expectation was 50%. Apart from the typically strong early fluctuations, there is a slow fluctuation of the cumulative average in the range of 54.5 ± 0.5%. The interval between trials with staring periods of usually 10s seems to have been on the order of minutes.

Using the direct method, Radin (2006) calculated the overall chance probability P for the datasets of each of the three psi effects. The probability was derived from the final cumulative, or overall z-score, which is obtained from the total number of trials and the overall hit rate by means of Equation (1). The overall z-scores were $z_{overall} = 6.49$, 9.07, and 16.44 for the dream psi experiments, the ganzfeld studies, and the tests of the sense of being stared at. Though not quoted by Radin, they are explicitly written down here because they will be needed in the following for a comparison of effect sizes. The chance probabilities P(z) associated with $z_{overall}$ are $P \approx 10^{-10}$, 10^{-19}, and 10^{-59}, respectively. The standard errors of the cumulative mean excess hit rates $h - p$ shown in Radin's plots, i.e. the r.m.s. deviations of the excess numbers of hits divided by $\sqrt{N}$, are on the order of 10%.

In all three types of studies, the psi effect per trial as expressed by the extra hit rate $h - p$ in excess of its MCE value seems rather stable after a sufficient number of trials. This is surprising because different laboratories and many different test persons took part in the experiments. There is of course no guarantee that the cumulative mean values will remain unchanged when new experiments are conducted in the future. For instance, admitting only subjects of proven talent or modifying some experimental details may in the long run have dramatic consequences.

Psi Effects Studied in Many-Trial Experiments of Fixed Length

A psi experiment with an MCE hit rate $p = 0.2$ is card guessing with ESP cards. These come in decks of 25 cards bearing one of the five symbols: circle, square, cross, star, or wavy lines. Since their invention in 1934, a large number of trials have been accumulated, ranging up to a million and more as a function of the experimental quality demanded for inclusion. Steinkamp (2005) recently gave a survey of reviews of the numerous studies, beginning with an attempt to determine the effect size per trial, e, as defined by Equation (2). Most probably, evaluations of e were based on 25 consecutive trials or multiples thereof. The overall result varied with experimental quality and other factors, ranging roughly from $e = 0.05$ to 0.02. Steinkamp also dealt with the dependence of the psi effect on psychological and physical conditions. It seems impossible in the case of ESP cards to define a more accurate mean value of the effect size.

Another psi effect with an MCE hit rate of 20% is studied in the ball drawing test. It was introduced and investigated by Ertel (2005) whose report was reviewed and supplemented with some unpublished data by Broderick (2007). In 231 experiments, one of 50 table tennis balls marked with 1, 2, 3, 4, or 5 in equal proportion was drawn from an opaque bag with the subject aiming for one of the five numbers. After recording the number and returning the ball, the bag was shaken to restore random mixing. The number of balls drawn in an experiment was either 240 or 360. The experiments were done in the years 1998 to 2002, with a different test person in all 231 cases. The total number of trials, draws in this study, was 71,760, while that of hits was 15,646, which is 9.0% above the mean chance expectation of 14,352. This corresponds to a hit rate h = 0.218 instead of its MCE value p = 0.2. From these numbers and Equation (1), the overall z-score was calculated to be $z_{overall} = 12.07$, which leads to $P \approx 10^{-32}$. Insertion of h and p into Equation (3) yields the overall effect size e = 0.045. A remarkable feature of Ertel's experiments is a conspicuously high frequency of near misses, e.g., 2 and 4 when 3 was the targeted and most often drawn number.

For a comparison of psi effects, it will be useful to know $\langle z_{sequ} \rangle$, i.e. the mean value of the z-scores, of the 231 sequences of trials in the ball drawing studies. From the overall z-score or the annual values of z, one computes $\langle z_{sequ} \rangle = 0.79$ or 0.80, respectively. In both cases, Stouffer's formula

$$z_{Stouffer} = \langle z_{sequ} \rangle \, \Omega^{1/2}, \tag{4}$$

is used inversely to compute $\langle z_{sequ} \rangle$ from $z_{Stouffer}$. The quantity substituting here for $z_{Stouffer}$ is the overall z-score or its annual value, respectively, while Ω is the number of experiments involved. Actually, $z_{Stouffer}$ as defined by Equation (4) is the mean overall z-score, if every z_{sequ} obeys the same Gaussian distribution regardless of N. This fact may help one to understand Equation (2) defining the z-score per trial. In the present computation, as in others, $z_{Stouffer}$ is replaced by or interpreted as $z_{overall}$ with hardly any loss in accuracy whenever z_{sequ} satisfies the condition just mentioned and Ω is very large.

Psi Effects Studied in Many-Trial Experiments of Variable Length

The PK studies of mind–matter interaction using binary random number generators (RNGs) or the fall of dice involve more than a billion bits (Radin & Nelson, 1989, 2000) and more than 2.5 million fallen dice (Radin & Ferrari, 1991). The experiments began in 1935 and 1959, respectively, after the invention of suitable machines. The abundance of trials is due to automatization which permitted the number N of bits or throws in an experiment to be varied over many orders of magnitude. The subjects intended to produce either a positive or a negative deviation of the number of hits from its MCE value. Again, the stated goal of these studies and their meta-analyses was to determine the effect size e, i.e. the PK effect per bit or fallen die. This quantity was thought not to depend on the number of trials in an experiment and thus to characterize the psi effect.

Interestingly, among numerous graphs the authors diplayed the distributions of psi-influenced z-scores. The number N of bits or fallen dice does not enter into these representations. Any MCE statistical distribution of hits is essentially Gaussian, if the hit rate is not too near either of its limits, zero, or unity, and if N is large enough. This was the case in all the experiments underlying the meta-analyses. In plotting the graphs, negative intentions (i.e. those aiming for fewer than the normal number of hits) were reinterpreted as positive intentions by changing the sign of the associated z-scores. The resulting plots were approximated by modified Gaussian distributions that were shifted in the direction of positive z. In addition, they were widened in comparison with the normal distribution, although this was not intended in any way. Radin and Nelson (2000) and Radin and Ferrari (1991) computed the overall z-score of the modified distribution from the z-scores of the experiments by means of Stouffer's formula. From this z-score, they calculated the overall chance probability of the dataset. A more complete treatment takes into account both the shift and the widening factor of the Gaussian distribution under the influence of psi (Helfrich, 2007). It results in a much smaller value of the overall chance probability. The treatment will be revisited below in the search for a relationship between shift and widening.

It is remarkable that Gaussian z-score distributions could be used, at least as plausible approximations, to describe the effect of psi, even though N varied by several orders of magnitude among the experiments. Their apparent indifference to N raises the suspicion that the z-scores of PK experiments with RNGs and falling dice are independent of the number N of trials in a sequence, or, in other words, that psi modifies the distribution of chance probabilities in the same way for all (large enough) N. This would clearly justify equating $z_{Stouffer}$ to $z_{overall}$ if Ω is large. The conjecture that $\langle z_{sequ} \rangle$ does not depend on the number N of random events in a PK experiment had been put forward earlier, mainly under two headings: the Intuitive Data Sorting model (IDS) of May, Radin, Hubbard, Humphrey, and Utts (1985) and the Decision Augmentation Theory (DAT) of May, Utts, and Spottiswoode (1995). These models explain PK in terms of precognition, which in the case of RNGs seems possible if the test person pushing the button enters a continuous predetermined random sequence of bits at an appropriate moment. The model runs into difficulties in the case of dice, where no sequence is entered, and is in conflict with systematic RNG studies of the PEAR (Princeton Engineering Anomalies Research) group (see below). Moreover, a random choice of bits brought about by quantum events such as nuclear decay, tunneling in diodes, or thermal scattering is incompatible with the idea of predetermination. Let me take a neutral position and simply assume that psi effects are goal-oriented and probably reach their goal on rather direct pathways which remain to be explored.

Binary RNGs. The original meta-analysis by Radin and Nelson (1989) was based on the z-scores of 597 experiments with binary RNGs. The authors defined as an experiment the largest possible aggregation of bits collected in a given report under a single directional intention. Pilot and confirmatory tests were treated as separate experiments. They did not distinguish between pure cases and others, probably rare, in which the intentions were intermingled. Also, the number of subjects following one another and the number of button pushes in an experiment were disregarded. (In this context, it is interesting to note that Radin and Nelson (1989) expressed results reported merely to be insignificant by the MCE z-score distribution but cut off at the limits $z = \pm 1.645$.) The mean shift $\langle z_{sequ} \rangle$ and the widening factor or standard deviation α as read from the plotted psi-modified Gaussian z-score distributions are $\langle z_{sequ} \rangle = 0.6$ and $\alpha = 1.5$. Analytically, the authors derived $\langle z_{sequ} \rangle = 0.645$ and $\alpha = 1.601$ from the data points underlying the modified distribution. The effect size was averaged over the experiments with a weight proportional to N and found to be about $(3 \pm 0.5) \times 10^{-4}$, with surprisingly little dependence on experimental quality and other factors. Even deleting 17% of the experiments as outliers, to achieve homogeneity of the e-values computed from the individual z-scores, did not make a marked difference. The Princeton Engineering Anomalies Research (PEAR) group, who processed more bits than any other psi laboratory, found the effect size to be on the order of $e = 1 \times 10^{-4}$ (Jahn, Dunne, Nelson, Dobyns, & Bradish, 1997, Jahn & Dunne, 2005).

In the updated meta-analysis of Radin and Nelson (2000), the 258 experiments of the PEAR group were collapsed into a single data point (or two), while new ones, 84 reported by 1987 and 92 reported after 1987, were added. For this dataset, the practical absence of a correlation between N and z of the RNG experiments was expressly stated. (The correlation of $\sqrt{N}$ and z was given as $r = -0.015$, $P = 0.36$, which is far below the level of statistical significance.) It can also be directly recognized from a graph in Schub's (2006) critique of Radin and Nelson (1989, 2000) that displays data points in a (z,logN) scatter diagram. Also in the update, the overall z-score was computed with Stouffer's method. The result, $z = 16.1$, corresponds to the chance probability $P(16.1) \approx 10^{-58}$. Particularly valuable for a comparison with other psi effects is a plot of the cumulative mean z-score per experiment as a function of the year of publication. Its final value for the period of 1959 to 1987 is $\langle z_{sequ} \rangle = 0.71 \pm 0.05$ at the top of a slight upward fluctuation following a bottom at 0.62. The post-1987 data are plotted separately, their overall mean value being $\langle z_{sequ} \rangle = 0.61 \pm 0.12$. Their standard error is large because of the relatively small number (92) of experiments. Since the RNG mean z-score is a reference point in the following, some additional details should be noted. Dividing the data of 380 selected RNG experiments into quartiles of 90 according to bit number N, Bösch, Steinkamp, and Boller (2006) found $\langle z_{sequ} \rangle$ to range from 1.05 (smallest N, z probably inflated by some gifted test persons) to 0.41 (largest N, but N > 10^9 omitted). Their overall mean was $\langle z_{sequ} \rangle = 0.67$. The value used for RNGs in the estimates below is $\langle z_{sequ} \rangle = 0.65$. The slight dependence of $\langle z_{sequ} \rangle$ on N would seem to be not in serious conflict with Radin and Nelson's (2000) finding that $\langle z_{sequ} \rangle$ is practically constant. The question of a general constancy of $\langle z \rangle$ that holds either for sequences or, in other situations, for single trials, will be an important topic in the following. The plot of cumulative mean z-scores can be regarded as a valid counterpart to the plots of cumulative hit rates of the first three psi effects since $h - p$ is easily converted into e, i.e. the mean z-score of a one-trial experiment.

It was pointed out above that an indifference of the z-score distribution to N implies, for the result of a single PK experiment, an N-independent distribution of chance probabilities. Let it be emphasized again that this is in sharp conflict with the more common assumption that the effect size or, equivalently, the psi-induced part of the hit rate, is independent of N. In the latter case the mean psi-induced z-score is expected to increase with $\sqrt{N}$, because the average difference of psi-affected and MCE hit numbers varies with N, while the standard error of the hit numbers is proportional to $\sqrt{N}$. The PEAR group adhered in general to the concept of a constant PK effect per bit. Dobyns and Nelson (1998) utilized the wealth of data at PEAR to compare this concept to the model of constant $\langle z_{sequ} \rangle$. In this comparison, a sequence generally consisted of the bits between button pushes, the bit number varying from 200 to 200,000. Another experimental variable was the number of bits in a 0.2s block (which at PEAR was called a trial). Apart from the regular 200-bit blocks, there were blocks of 20 and 2,000 bits. It turned out that the data agree better with a constant effect size than with a constant $\langle z \rangle$ per sequence of bits.

Recently, it was proposed that there may be room for both models, that of constant e as well as that of constant $\langle z_{sequ} \rangle$ (Helfrich, 2007). The constant-$\langle z_{sequ} \rangle$ model apparently applies to isolated experiments, i.e. those which in motivation, time, and probably other factors are well separated from others of the same kind. A constant effect per bit, independent of N, appears to prevail whenever a particular experiment is serial or, more aptly, "embedded" in a series of similar experiments, which seems to have been the typical situation at PEAR. The 515 experiments underlying the updated meta-analysis of Radin and Nelson (2000) were, of course, not distinguished with respect to the degree of isolation. Confidence that they were sufficiently isolated can be based only on the Gaussian z-score distribution and the small number of experiments in most of the 216 original publications.

The conflict between the two models is confusing, and there is a general tendency to fall back on the concept of an N-independent effect per bit because this hypothesis appears more physical. Acceptance of an N-independent z-score may be made easier by a recent observation of Jahn and Dobyns (2007). They pointed out that a $\langle z_{sequ} \rangle$ can be independent of N not only if the effect size equals $e(N) = \langle z_{sequ} \rangle / \sqrt{N}$ for all bits in a sequence of N bits, but also if it equals $e(n) = \frac{1}{2} \langle z_{sequ} \rangle / \sqrt{n}$ at any N, with n being the running number of bits. The equivalence is a consequence of the identity

$$(1/2N) \int_0^N (1/n^{1/2}) dn = 1/\sqrt{N} \qquad (5)$$

Utilizing previous PEAR data, the authors tried to find out which of the two variants of the constant-$\langle z_{sequ} \rangle$ model is the better one, without reaching a conclusion. No such check has as yet been made exclusively for isolated experiments.

Also at PEAR, Ibison (1998) and later on Dobyns, Dunne, Jahn, and Nelson (2004) examined the constancy of the effect size in a dramatic fashion. In otherwise unchanged experiments, they increased the bit rate from the regular 200 per block to 10,000 times as many. A block of random bits was generated, in the usual manner, within an interval of 0.2s in every period of 0.9s. Low speed, i.e. the usual bit rate, was obtained by taking only one of every 10,000 bits. Both studies involved many test persons. In Ibison's study, 70 series of 1,000 blocks were recorded with each intention (high, low, baseline). For a given intention, high-speed and low-speed sequences alternated at random from block to block but were sorted afterward according to speed, thus representing two experiments in the evaluation. The focus was on the difference effect, i.e. the difference of the z-scores with positive and negative intention divided by $\sqrt{2}$. At low speed, the overall difference z-score was 1.2967. This is insignificant, but happens to agree very well with what may be expected on the basis of Radin and Nelson's meta-analysis (1989) and rather well with the large body of experiments by PEAR that employed 200-bit blocks. In fact, inserting $N = 7 \times 10^6$ into Equation (2), one obtains from the overall difference z-score divided by $\sqrt{2}$ the effect size $e = 3.5 \times 10^{-4}$. However, this inference is of little weight and, in particular, does not prove that the experiments were of the "embedded" type, because of the insignificance and, correspondingly, large standard error of the z-score. At high speed, the difference z-score was −3.7391, which is significant ($P = 1 \times 10^{-4}$), but larger than the low-speed result only by a factor near three while a factor of 100 may be expected on the basis of Equation (1). In other words, the effect size as computed from the overall z-score decreased by a factor of roughly 30. Even more surprisingly, the sign of the overall z-score at high speed was contrary to intention. Apart from essentially confirming Ibison's results in a total of 149 similar experiments, Dobyns et al. did 39 experiments solely at high speed, noting no significant difference in effect size from the high-speed part of the experiments mixing the speeds. The findings of Ibison (1998) and Dobyns et al. (2004) agree with neither the constant-$\langle z_{sequ} \rangle$ nor the constant-e model. However, they suggest that the PK effect per bit tends to be suppressed when otherwise the absolute value of the z-score would become "too large". One may wonder whether the two studies, if taken as two times three single experiments, could be of the isolated type. In this case mean difference z-scores on the order of z = 1 would be within the expected range, but mean difference z-scores on the order of z = 4, as found at high speed, are outside. The change in the sign of z cannot be explained by any existing model. These studies suffer from large errors because, like many specific studies, they involve relatively small numbers of experiments.

Falling dice. Radin and Ferrari's meta-analysis of the PK effect on falling dice was based on 148 experiments. They obtained for the effect size e = 0.012 ± 0.006 the large uncertainty arising from a considerable dependence on experimental quality and other factors. The mean shift and the widening factor as read from the drawn Gaussian distribution of psi-affected z-scores of dice experiments are $\langle z_{sequ} \rangle = 1.5$ and $\alpha = 2.5$. The overall z-score was computed to be $z_{Stouffer} = 18.2$, resulting in a chance probability of 10^{-74}. There is no plot of the cumulative mean z-score per experiment over time and no check of the correlation between z and N in the meta-analysis of the dice experiments.

The PK effect on falling dice has not only been investigated less often and mostly prior to the advent of RNGs, it is also more prone to pitfalls than that on RNGs. Perhaps the greatest problem is the fact that even without PK the hit rates of the six faces of a die usually are not identical. For instance, some material is lost if the numbers are marked by small scoops, which favors the six-face to end up on top. Radin and Ferrari (1991) showed in their Figure 6 that the effect of the asymmetry is nearly as large as the PK effect. To avoid mechanical effects, one has to admit only balanced experiments in which all six faces are equally often the target of intention. This reduces the number of experiments underlying the meta-analysis from 148 to 69. With this restricted dataset, Radin and Ferrari computed $z_{Stouffer} = 7.617$, which correponds to a chance probability of 10^{-14}. Reverse use of Stouffer's formula (4) led from this number to the average z-score $\langle z_{sequ} \rangle = 0.917$ (±0.1) of the 69 experiments. The situation may be even more complicated because Figure 6 of Radin and Ferrari (1991) also seems to show that certain die faces, in particular 6, 2, and 1, are preferred in PK studies even after correcting for the effect of mechanical asymmetry.

There is another serious problem with the dice data because, as a rule, several dice are thrown at once, the upper limit apparently being 96 (Radin, 2006). Both Radin and much earlier Rhine (1972) provide data suggesting that the PK effect per die is independent of the number of dice in a throw. Let me adopt this independence as an assumption, although it seems unlikely to hold for extremely large numbers of dice. Otherwise, a single throw of a sufficient number of dice could generate a mean z-score of any value desired.

Tentative Conclusions Drawn from Comparing Mean Values of Psi Effects

The most important data of the psi studies considered here are compiled in Table 1 for convenient comparison. Inspection of the data suggests examination of three types of possible relationships between mean values of psi effects. First, the mean z-scores of the one-trial ESP experiments are not much smaller than the mean z-score of many-trial PK experiments on RNGs. How close is this similarity at best and what could it mean? Second, the widening factor α of the shifted and widened z-score distributions which ideally is identical to its standard deviation may be regarded as a mean value. Can a simple explanation be put forward for the widening? Is there a relationship between α and $\langle z_{sequ} \rangle$, i.e. widening and shift, that holds for binary RNGs, falling dice, and, perhaps, other PK or ESP effects? Third, the mean shift of the z-scores found with binary RNGs is smaller than those found with falling dice and ball drawing. These experiments differ by the number of targets that are equivalent in the absence of psi or, for short, by their multiplicity m. Can a plausible formula be proposed that relates the mean z-score at any m to its particularly trustworthy value at $m = 2$? To find such a relationship was the goal of a preliminary publication (Helfrich, 2008). It may be difficult to achieve, even when more data become available, if indeed psi effects with $m > 2$ discriminate between choices that without psi are equivalent. Finally, an update complementing an earlier article (Helfrich, 2007) is given of psi-induced switching and its possible role in mind–neuron interaction.

A Comparison of One-Trial Experiments in ESP to Many-Trial Experiments in PK

A glance at Table 1 reveals that the mean z-scores per trial, i.e. $\langle z_{trial} \rangle = z_{overall}/\sqrt{N}$, of the studies of dream psi, ganzfeld psi, and the sense of being stared at, are less than an order of magnitude below the mean z-score $\langle z_{sequ} \rangle$ of (presumably) isolated RNG experiments usually consisting of a large number of trials. As already mentioned, the number of trials, N, has different meanings in the two cases. In the first three studies it is the total number collected in the course of years, while in the fourth study N refers to a single experiment. The extreme effect sizes of the one-trial experiments may be due to a high degree of isolation of the trials. The dream psi experiments seem to be particularly well isolated from each other since no more than one picture was “sent” in a night. The increase of the hit rate from 50% to 59.1% represents the maximum effect

TABLE 1
Compilation of Data on PK and ESP Experiments

Type of Experiment	Dream	Ganzfeld	Staring	Ball Test	ESP Cards	Bits	Dice, All	Dice, Balanced
Trials	1270	3145	33357	71760	$>2\cdot10^6$	$>1.4\cdot10^9$	$>2\cdot10^6$	?
Sequences	na	na	na	231	na	515	148	69
Hit rates (in percent)	59.1/50	32/25	54.5/50	21.8/20	nc /20	nc	nc	nc
$z_{overall}$	6.49	9.07	16.44	12.07	nc	nc	nc	nc
$\langle z_{trial}\rangle = e$ isolated	0.182	0.162	0.090	na	na	na	na	na
$\langle z_{trial}\rangle = e$ embedded	na	na	na	0.045 in sequ	0.02–0.05	$3\cdot10^{-4}\pm 5\cdot10^{-5}$	0.0122 ± 0.0062	0.00861 ± 0.00110
$\langle z_{sequ}\rangle$	na	na	na	0.79	na	0.65, (0.41–1.05)	1.5	0.917
$z_{Stouffer}$	na	na	na	na	na	16.1	18.2	7.617
Widening factor α	na	na	na	?	?	1.5	2.5	?
Reference	Radin (2006)	Radin (2006)	Radin (2006)	Ertel (2005)	Steinkamp (2005)	Radin & Nelson (1989, 2000)	Radin & Ferrari (1991)	Radin & Ferrari (1991)

na, not applicable. nc, not calculated. Trials means number of individual random events. Sequences means number of (presumably) isolated sequences. Hit rates are overall psi-influenced values (above) and MCE values (below). (See also main text.)
$z_{overall}$ is the z-score taken over all one-trial experiments for a given psi effect.
$\langle z_{trial}\rangle = e$ is the effect size or effect per trial, which equals $z_{overall}/\sqrt{}$(total number of trials) in the case of isolated trials, but is an average of $z_{sequ}/\sqrt{}$(number of trials in a sequence), in the case of embedded many-trial experiments.
$\langle z_{sequ}\rangle$ refers to (presumably) isolated many-trial experiments only. The values of $\langle z_{sequ}\rangle$ in parentheses are taken from Bösch, Steinkamp, & Boller (2006).
$z_{Stouffer}$ represents $\langle z_{sequ}\rangle \times \sqrt{}$ (number of presumably isolated sequences).

size in Table 1. The ganzfeld experiments with an increase from 25% to 32% are almost equally impressive, but more difficult to compare as they are of type m = 4. Here the time taken for sending and receiving a single picture, including preparations, seems to have been an hour or two. Experiments testing the sense of being stared at, which permit repetition probably in a matter of minutes, resulted in another high hit rate, 54.5% instead of normally 50%, but the excess above 50% is only half as large as in the dream psi experiments.

There is still a gap by a factor of 3.6 between $\langle z_{sequ}\rangle = 0.65$ as found in psi experiments with binary RNGs and $\langle z_{trial}\rangle = 0.182$ as found in dream psi studies. The number of equivalent choices is m = 2 in both cases, but one effect is ascribed to PK and the other to ESP. Let me try to reduce the gap further with the following argument based on Equation (5). Replacing the integral in this equation by the sum of the discrete values of $1/n^{-1/2}$ over all n and keeping only the first term (n = 1) in an attempt to realistically describe the case N = 1, leads to ½ instead of 1 on the left-hand side of this equation, while 1 is obtained on the unaffected right-hand side. This suggests $e_{N=1} = ½\langle z_{sequ}\rangle$, so that $\langle z_{trial}\rangle$ of the dream psi studies would be expected to be half as large as $\langle z_{sequ}\rangle$ of RNG psi experiments at large bit numbers ($N \geq 20$) where the difference between integral and sum is negligible. The gap thus shrinks to the factor 1.8, and the two mean z-scores of dream psi, one actual and the other "theoretically" predicted, may be regarded as "practically equal", considering the uncertainties and fluctuations mentioned elsewhere in the present article. Of course, this idea is speculative in more than one respect. For instance, replacing for N = 1 the integral in Equation (5) by the first term of a sum seems crude, but it might be at least a step toward a more accurate treatment. Unlike the mean z-score of sequences, that of one-trial experiments is limited by the maximum effect size which is e = 1 in the case m = 2 (see below). The closeness of $\langle z_{trial}\rangle = 0.182$ to that limit might be a further factor explaining its small value. In the tests of the sense of being stared at, another binary choice, $\langle z_{trial}\rangle$ is half as large as in dream psi. This could be due to the much higher frequency of the guesses which possibly impairs their isolation. A mean z-score proportional to the square root of the time spent for achieving it has been proposed by Nelson (2006) in his theory of a time-normalized yield in psi experiments.

If the considerations of the last two paragraphs are correct, three types of mean z-scores may have to be distinguished in psi experiments. The first is that of a single trial isolated from all other one-trial and many-trial experiments of the same kind. The second is the mean z-score of a rapid sequence of (many) trials which is isolated from all similar experiments including single trials. The third is that of a trial, either single or part of a sequence, embedded in an environment of similar experiments.

The fact that after the correction $\langle z_{trial}\rangle$ of dream psi experiments and $\langle z_{sequ}\rangle$ of RNG experiments, both of multiplicity m = 2 appear to be equal or nearly so suggests two tentative conclusions: First, psi is (almost) equally effective in PK and ESP, and, second, the mean z-score of an isolated experiment depends little on the number of trials involved all the way down to N = 1. The conclusions are interdependent, i.e. both of them are either right or wrong. The mean z-score for m = 2 seems to lie in the vicinity of 0.5 in one go or, in psychological terms, upon one impulse of motivation. This is when there is no interference with "nearby" similar experiments which appears to reduce the psi effect. Conversely, one may wonder if several impulses can act in long sequences, especially those performed by a series of subjects, thus augmenting the z-score.

Looking for a Relationship between Psi-Induced Shift and Widening of the MCE-Gaussian Distribution

Let it first be shown that the widened Gaussian distribution can be understood as the product of two Gaussian distributions. The normalized versions are being used whose integrals are unity. One of them is the MCE distribution without PK influence,

$$(2\pi)^{-1/2}\exp(-z^2/2). \qquad (6)$$

(For convenience, the subcript of z_{sequ} is dropped here and in some of the following formulas.) The other is due to a newly introduced scatter of the PK effect per bit or falling die,

$$(2\pi)^{-1/2}\, s_{PK}{}^{-1} \exp[-(\zeta^2/2s_{PK}{}^2)], \qquad (7)$$

where s_{PK} is the standard deviation of the PK effect which, like ζ, is measured in units of z. (The mean shift of z can be ignored in these considerations.) Integrating the normalized product function $(4\pi s_{PK})^{-1} \exp[-(z+\zeta)^2/2 - \zeta^2/(2s_{PK}{}^2)]$ over ζ for a given z leads to a widened Gaussian distribution as a function of z

$$[2\pi(1 + s_{PK}{}^2)]^{-1/2} \exp[-z^2/2(1 + s_{PK}{}^2)]. \qquad (8)$$

It follows from the final form that the widening factor α of the Gaussian distribution modified by psi obeys the relationship

$$\alpha^2 = 1 + s_{PK}{}^2. \qquad (9)$$

Therefore, the widening can be interpreted as the consequence of a Gaussian scattering of the PK effect. To avoid averaging back to the mean effect size (zero in the present example), the fluctuations of the effect size must be slow as compared to the duration of an experiment. For binary RNGs, a different model was proposed by Pallikari (2004, 2008) who attributes the widening to a tendency of equal bits to agglutinate in the presence of psi. However, at least in the PEAR data no such effect was noted (Nelson, 2008).

Inspecting the z-score distributions of the PK effects on binary RNGs and dice immediately shows that the increase in width, $\alpha - 1$, roughly equals the mean shift $\langle z_{sequ} \rangle$ in both cases. As a proportionality of these two quantities seems mathematically unsound, it is preferable to use the equally simple equation

$$s_{PK} = \gamma \,|\, \langle z_{sequ} \rangle \,|\,, \qquad (10)$$

assumed to hold for RNGs and dice and, it is hoped, any number of equivalent choices. This "geometrical" model gives $\gamma = 1.72$ when the values for binary RNGs, $\langle z_{sequ} \rangle = 0.65$ and $\alpha = 1.5$, are inserted into Equations (9) and (10) and s_{PK} is eliminated between these equations. If proportionality applies, the same value of γ should be obtained with dice. Inserting $\alpha = 2.5$ and $\langle z_{sequ} \rangle = 1.5$, the value found in the meta-analysis of all 148 dice experiments, results in $\gamma = 1.53$. Using instead the mean z-score of the 69 balanced experiments, $\langle z_{sequ} \rangle = 0.917$, yields $\gamma = 2.50$. Both values of γ for dice differ from that for RNGs, one being smaller and the other larger, but with $\langle z_{sequ} \rangle = 1.5$ agreement is clearly better than with $\langle z_{sequ} \rangle = 0.917$.

An alternative model may start from the formula for the mean chance probability of a single experiment, i.e. for the factor φ by which the overall value of P_2, another kind of chance probability, is reduced on average by an additional experiment under the influence of psi (Helfrich, 2007). Taking account of the fact that the number of possible z-scores per standard deviation is proportional to α, one has for the shifted and widened distribution the reduction factor

$$\varphi = \alpha \times \exp[-(\langle z^2 \rangle - 1)/2] , \tag{11}$$

while the total chance probability for the result of Ω experiments is $P_2 = \varphi^{\Omega}$. Use of $\langle z^2 \rangle = \alpha^2 + \langle z \rangle^2$, which follows from integrating over z the product of z^2 and the normalized Gaussian distribution function of Equation (6), transforms Equation (11) into

$$\varphi = \exp[\ln\alpha - (\alpha^2 - 1 + \langle z \rangle^2)/2] \tag{12}$$

Incidentally, the exponent of the exponential function can be interpreted as minus the free energy which it costs to move a single particle from the MCE to the modified Gaussian distribution. As in the exponent of a Boltzmann factor, the energy is understood to be divided by (or given in units of) kT, where k is Boltzmann's constant and T absolute temperature. The potential is entropic, reflecting the number of ways in which a particular z-score can be attained. Accordingly, the internal energy of the hypothetical one-particle ideal gas is purely kinetic. It can be ignored, being conserved under shift and widening.

The digression into physics is not needed to realize that the exponent in Equation (12) may be split into the terms $\ln\alpha - (\alpha^2 - 1)/2$ and $-\langle z \rangle^2/2$. This allows separating the contributions of shift and widening to the reduction factor φ of the modified Gaussian distribution. For RNGs with $\alpha = 1.5$ and $\langle z \rangle = 0.65$ the terms are -0.22 and -0.21, respectively. Therefore, in the case of RNGs the reduction factor due to widening practically equals the reduction factor due to shift, although widening was not intended in the experiments. A model relating mean shift and widening may now be based on the expectation that this equality applies as well to PK effects of other multiplicities. For falling dice with $\alpha = 2.5$ and $\langle z \rangle = 1.5$ (or 0.917) the two terms in the exponent of Equation (12) are −1.71 and −1.125 (or −0.42), respectively. Evidently, for both values of $\langle z \rangle$ the "free energy" of widening is distinctly larger than that of shifting. The discrepancy is smaller with $\langle z \rangle = 1.5$ than with $\langle z \rangle = 0.917$. On the whole, it seems that the geometrical model is the more attractive one and that the larger mean shift is the "better" result, despite the fact that it includes unbalanced studies. However, the dice data are not sufficient in number and quality to rely on them. As yet, it is not possible to decide which of the two models proposed is the correct one or if both of them fail.

The detailed data of Ertel's (2005) experiments, as yet unpublished, should reveal whether widening also occurs in the ball drawing test. Interestingly, the Global Consciousness Project initiated by Nelson et al. (1996) uses the widening of the z-score distributions in a network of binary RNGs to monitor events that arouse worldwide emotion.

Looking for a Dependence of the Mean z-Score on the Number of Equivalent Choices

The question of a possible dependence of $\langle z \rangle$ on the multiplicity m will be discussed largely in terms of many-trial experiments, but the results can always be transferred to one-trial experiments by putting $N = 1$. The simplest and most extensively studied psi effect is the electronic version of coin throwing carried out with binary RNGs. Also, psi effects of $m = 2$ seem to be immune to the psychological perturbations which apparently plague psi effects of higher multiplicities. Therefore, they are a suitable reference point in dealing with psi effects of $m > 2$. The foremost question is whether the mean z-score per isolated one- or many-trial experiment varies with the number m of choices that are equivalent without psi. To begin with, let me propose two different mathematically straightforward models. The first one is based on the trivial assumption that $\langle z \rangle$ is independent of m. If the problem of widening is left aside, the model means in physical terms that the free energy of shifting the Gaussian distribution of z-scores is the same for all m. This "constant-$\langle z \rangle$" model seems to fail when applied to the fall of dice with $m = 6$ and the ball test with $m = 5$. In both cases, the mean z-scores per experiment are larger than those of binary RNGs. However, the difference is small for $m = 5$ where $\langle z_{sequ} \rangle$ is 0.79 instead of 0.65. A direct confirmation of the constant-$\langle z \rangle$ model in the case of one-trial experiments could be the near equality of the effect sizes in the dream and ganzfeld psi studies where $m = 2$ and 4, respectively.

The second model may be called the "generalized binary model." It is based on the idea that the MCE frequency of misses is reduced by a psi-induced partial conversion of a miss into a hit. The efficiency of conversion is assumed to be equal for all misses and not to depend on multiplicity. This is an extension of the binary model to every pair of a miss and the target. Accordingly, the number of hits in excess of its MCE value is taken to be e(1 − p)N, so that the mean z-score may be written as

$$\langle z \rangle = e(1-p)N/[p(1-p)N]^{1/2} \quad (13)$$

If a probable small difference between the mean z-scores for single-trial and many-trial experiments is disregarded or comparison is restricted to one type of experiments, the mean z-scores depend solely on the multiplicity $m = 1/p$ of the experiments. This will now be indicated by the subscript m. Substituting in Equation (13) $\langle z \rangle_2$ for $e\sqrt{N}$ and cancelling $\sqrt{N}$ leads to

$$\langle z \rangle_m = \langle z \rangle_2 (1-p)/[p(1-p)]^{1/2} \quad (14)$$

Note that the denominator in Equation (14) tends toward zero with decreasing p. Straightforward manipulations transform this into the simple form

$$\langle z \rangle_m = (m-1)^{1/2} \langle z \rangle_2. \quad (15)$$

This relationship was recently proposed (Helfrich, 2008) to explain the large difference of the shift of the z-score distribution of falling dice from that of binary RNGs. Incidentally, the generalized binary effect bears a similarity to the above-mentioned indifference of the PK effect per die on the number of dice in a throw.

The ratio of the shifts $\langle z \rangle_6/\langle z \rangle_2 = \sqrt{5} = 2.34$ predicted by Equation (15) is indeed close to the experimental value 1.5/0.65 = 2.31 obtained with the mean shift of the dice experiments if all of them are included. Equation (15) fails when only the balanced dice experiments are considered, which reduces the ratio of the shifts to 0.917/0.65 = 1.41. It also fails when applied to Ertel's (2005) ball drawing test, as it predicts $\langle z \rangle_5/\langle z \rangle_2 = 2$, while the experimental value is $\langle z \rangle_5/\langle z \rangle_2 = 0.79/0.65 = 1.22$. The latter could, perhaps, be raised markedly if the psi-induced near misses reported by Ertel can be counted among the hits. A questionable feature of Equation (15) is the predicted divergence of $\langle z \rangle_m$ with m, suggesting a method to achieve high levels of significance in a single psi experiment.

It is tempting to speculate that the two simple formulas just introduced might be useful as the lower limit (the constant-$\langle z \rangle$ model) and the upper limit (the generalized binary model) for $\langle z_{sequ} \rangle$ or $\langle z_{trial} \rangle$ in a psi experiment of multiplicity $m > 2$. If the constant-entropy model is regarded as the basic one, "proximity" or "similarity" or other psychological effects could increase $\langle z \rangle_m$ up to the limit prescribed by the generalized binary model. Both the near misses in Ertel's ball drawing tests and a preference for certain die faces in the fall of dice are possible examples. In one case it seems to be numerical proximity, in the other it might be a subconscious predilection for certain numbers. Such effects could result, whenever $m > 2$, in a deformation of the distribution of hits which depends on the specifics of the experiment.

Switching Properties of PK for Multiplicity m = 2

How the mind may act on the neuron via the neuron's synapses has been discussed elsewhere (Helfrich, 2007). The general problem of psi-induced switching in the case $m = 2$ may be worth being treated again, mainly to include one-trial experiments and the effect of widening. The limiting hit rates of statistical psi effects are unity and zero, i.e. a hit every time or no hit at all. Because of Equation (3) and $m = 1/p$, the associated effect sizes are $e = \sqrt{(m - 1)}$ and $e = -1/\sqrt{(m - 1)}$, which for $m = 2$ take the values $e = 1$ and $e = -1$, respectively. If the hit rate is 50% in the absence of psi and 60% in its presence, as in dream psi experiments, the surplus rate of hits over misses is 20% of the normal hit rate, but 50% of the psi-reduced rate of misses. This is a substantial effect and not too many single-trial experiments are needed to achieve statistical significance. Of particular interest seems to be the case of many-trial experiments at large enough N where a Gaussian distribution of z-scores exists and is shifted by PK. Let me assume that a critical z-score separating on- and off-states is located at $z = 1.96$, so that without psi the chance probability for $z \geq 1.96$ is $P(1.96) = 2.5\%$, a typical value for the limit of significance. The state $z \geq 1.96$ is regarded as the on-state, while lower z-scores belong to the off-state. In the presence of psi with a non-fluctuating $z_{psi} = 0.65$, the chance probability for $z \geq 1.96$ becomes $P(1.96 - 0.65) = P(1.31) = 9.5\%$. If widening is taken into account, the argument of P in the last equation has to be divided by the widening factor $\alpha = 1.5$, which leads to $P(0.87) = 19.1\%$. These ideas may now be applied to mind–neuron interaction. Here the z-score represents an electric potential divided by its standard deviation from its MCE value. The critical value of z is taken to represent the threshold potential which, when it is exceeded, causes the neuron to fire a new action potential into its axon. The circa 10,000 synapses of the neuron are assumed to act like 50:50 probabilistic switches when an action potential arrives through one of the presynaptic axons. Together, they thus mimic a binary RNG emitting, instead of bits, either a small amount of electric charge or no charge. A complete theory has to include the dynamics of the neuron potential, replaced here by potentials that are static during each switching period (of a few ms). Taking into account the dynamics probably lowers the quality of switching, while the fluctuations of the z-score might improve it. An increase of the probability of firing an action potential from 2.5% to 19.1% may appear too small for sufficiently reliable switching. Several switching processes in series or in parallel may be required for a useful effect. On the other hand, probability differences near 100% seem unreasonable. After all, one has to expect that psi allows influence but no control, at least in everyday situations. In particular, while persons may be able to communicate to some extent by means of psi, they should not be able to enslave each other in this way. Similar anthropic arguments with regard to $\langle z_{sequ} \rangle$ were put forward in the previous article.

Concluding Remarks

The search for relationships between mean values yielded three intriguing findings. First, the mean z-score at m = 2, i.e.with two equivalent choices, is nearly indifferent to the number of trials in isolated experiments down to N = 1. Second, it is equal or nearly so in PK and ESP. Third, the widening of the z-score distributions of many-trial experiments can be explained in terms of a Gaussian scattering of the effect size. The three relationships require, of course, further examination before they can be considered proved. In particular, the Gaussian shape assumed for the psi-modified z-score distributions of PK experiments with RNGs and dice is so far only an approximation to rather rough experimental distributions. Two models were proposed that might give the widening factor as a function of the shift at any m if the relationship is known for m = 2. One of them may be correct, but cannot yet be definitively checked because of problems with the data on falling dice and unavailability of the data in the case of ball drawing. Moreover, two attempts to relate the mean z-score at arbitrary m to that at m = 2 seem to fail. It was argued that this may be due to a discrimination, under the influence of psi, between targets of equal MCE hit rate.

A basic equality of the mean z-scores at m = 2 in isolated PK and ESP experiments, and only a weak dependence of them on the number of trials in an experiment, would be in accordance with an often-noted far-reaching indifference of psi experiments to distances in space and time and, in the case of PK, to the type of binary RNG. The order of magnitude of the mean z-scores, which seems to emerge in such experiments, appears compatible with a possible role of psi in mind–neuron interaction and extrasensory communication. The conjecture that these two phenomena (of enormous philosophical and practical implications) do exist may be no more than fantasy, but it is supported by experimental results in statistical parapsychology.

Acknowledgments

I am particularly grateful for support to the following three persons. Eberhard Bauer generously supplied copies of articles and pointed out important recent literature. He and Walter von Lucadou gave me the opportunity to present the contents of this article at the October 2009 meeting of the Wissenschaftliche Gesellschaft für Parapsychologie in Offenburg, Germany. Roger Nelson encouraged me to write this article, read a preliminary version of the manuscript, and made some cautionary comments (which I did not always heed).

References

Bösch, H., Steinkamp, F., & Boller, E. (2006). In the eye of the beholder: Reply to Wilson and Standish (2006) and Radin, Nelson, Dobyns, and Houtkooper (2006). *Psychological Bulletin, 132,* 533–537.

Broderick, D. (2007). *Outside the Gates of Science.* New York: Thunder's Mouth Press.

Dobyns, Y. H., Dunne, B. J., Jahn, R. G., & Nelson, R. D. (2004). The MegaREG Experiment: Replication and interpretation. *Journal of Scientific Exploration, 18,* 369–397.

Dobyns, Y. H., & Nelson, R. D. (1998). Empirical evidence against decision augmentation theory. *Journal of Scientific Exploration, 12,* 231–258.

Ertel, S. (2005). The ball drawing test: Psi from untrodden ground. In: M. A. Thalbourne, & L. Storm (Eds.), *Parapsychology in the Twenty-First Century* (pp. 90–123). Jefferson, NC: McFarland & Company.

Helfrich, W. (2007). Is the psychokinetic effect as found with random number generators suitable to account for mind–brain interaction? *Journal of Scientific Exploration, 21,* 689–705.

Helfrich, W. (2008). Is there a relationship between the psychokinetic effects on the fall of dice and on random number generators? *Proceedings of The Parapsychological Association 51st Annual Convention & The Society for Psychical Research 32nd Annual Convention* (pp. 344–347).

Ibison, M. (1998). Evidence that anomalous statistical influence depends on the details of the random process. *Journal of Scientific Exploration, 12,* 407–423.

Jahn, R. G., & Dobyns, Y. H. (2007). Dependence of anomalous REG performance on run length. *Journal of Scientific Eploration, 21,* 449–472.

Jahn, R. G., & Dunne, B. J. (2005). The PEAR proposition. *Journal of Scientific Exploration, 19,* 195–245.

Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H., & Bradish, G. J. (1997). Correlations of random binary sequences with pre-stated operator intention: A review of a 12-year program. *Journal of Scientific Exploration, 11,* 345–367.

May, E. C., Radin, D. I., Hubbard, G. S., Humphrey, G. S., & Utts, J. M. (1985). Psi experiments with random number generators: An informational model. *Proceedings of The Parapsychological Association 28th Annual Convention*, vol. 1, 237–266.

May, E. C., Utts, J. M., & Spottiswoode, S. J. P. (1995). Decision augmentation theory: Towards a model of anomalous mental phenomena. *Journal of Parapsychology, 59,* 195–220.

Nelson, R. D. (2006). Time-normalized yield: A natural unit for effect size in anomalies experiments, *Journal of Scientific Exploration, 20,* 177–199.

Nelson, R. D. (2008). Personal communication.

Nelson, R. D., Bradish, G. J., Dobyns, Y. N., Dunne, B. J., & Jahn, R. G. (1996). FieldREG anomalies in group situations. *Journal of Scientific Exploration, 10,* 111–142.

Pallikari, F. (2004). On the false hypothesis of psi-mediated shift of statistical average in tests with random event generators. *Proceedings of The Parapsychological Association 51st Annual Convention*, pp. 157–171.

Pallikari, F. (2008). In search of the universal character of micro-psychokinesis. *Proceedings of The Society for Psychical Research 32nd Annual Convention*, pp. 382–383.

Radin, D. I. (2004). On the sense of being stared at: An analysis and pilot replication. *Journal of the Society for Psychical Research, 68,* 246–253.

Radin, D. I. (2006). *Entangled Minds*. New York: Paraview Pocket Books.

Radin, D. I., & Ferrari, D. C. (1991). Effects of consciousness on the fall of dice: A meta-analysis. *Journal of Scientific Exploration, 5,* 61–83.

Radin, D. I., & Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics, 19,* 1499–1514.

Radin, D. I., & Nelson, R. D. (2000). *Meta-Analysis of Mind–Matter Interaction Experiments: 1959–2000.* Saratoga, CA: Boundary Institute. [Also in: (2003) Research on mind–matter interaction: Individual intention. In: W. B. Jones, & C. C. Crawford (Eds.), *Healing, Intention and Energy Medicine: Science, Research Methods and Clinical Implications* (pp. 39–48), Edinburgh, UK: Churchill Livingstone/Elsevier.]

Rhine, L. E. (1972). *Mind over Matter*. New York: Collier Books.

Schub, M. H. (2006). A critique of the parapsychological random number generator meta-analyses of Radin and Nelson. *Journal of Scientific Exploration, 20,* 402–419.

Sheldrake, R. (2003). *The Sense of Being Stared At*. New York: Crown Publishers.

Sherwood, S. J., & Roe, C. A. (2003). A review of dream ESP studies conducted since the Maimonides dream ESP studies. In J. Alcock, J. Burns, & A. Freeman (Eds.), *Psi Wars: Getting to Grips with the Paranormal*, Thoverton, UK: Imprint Academic.

Steinkamp, F. (2005). Forced-choice ESP experiments: Their past and their future. In: M. A. Thalbourne, & L. Storm (Eds.), *Parapsychology in the Twenty-First Century* (pp. 124–163), Jefferson, NC: McFarland & Company.