Journal of Scientific Exploration, Vol. 31, No. 3, pp. 387–418, 2017 0892-3310/17

RESEARCH ARTICLE

Statistical Parapsychology as Seen by an Applied Physicist

WOLFGANG HELFRICH

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

Submitted August 12, 2016; Accepted March 3,2017; Published September 15, 2017

Abstract—An attempt is made to recognize a system behind the statistical psi effects that are evaluated in terms of hit rates. For this purpose, I formulate five rules that appear to apply at least to studies of good quality with the most common chance hit rates $p = ½$ or ¼. A problem in the evaluation of the results arises from the fact that the hit rate h cannot be smaller than 0 or larger than 1. This implies that the z-scores of an experiment, i.e. the ratio of deviation to standard deviation, and their mean values $< z >$ can be limited as well. The true effect size should in principle be unbounded, but its standard definition by $< z >$ may be expected to fail whenever h is near one of its boundaries. In order to deal with such a situation, most likely if an experiment consists of a single decision between hit or miss, an effect size is needed that is unlimited but for $(h - p) \rightarrow 0$ merges with $< z >$. Two such effect sizes are derived here from models of psi effects. Moreover, on the basis of a sixth rule, as yet preliminary, the scattering of the effect size, a common but little-explored phenomenon, and its possible consequences for the hit rate are dealt with. The comparison of the ratio of the $< z >$-scores of two extensively investigated psi effects with that of the corresponding conjectured true effect sizes helps to decide between the models. Another such comparison may suggest insufficient separation of (ganzfeld-psi) experiments.

Introduction

On one hand, parapsychology deals with rare phenomena that very likely are anomalous and in general are not repeatable. They lack the reproducibility characteristic of the natural sciences. On the other hand, roughly since the middle of the last century there have been numerous investigations of psi effects of a very different, nearly opposite kind: The result of a single experiment remains within random noise, but the statistical analysis of a large number of equal or similar experiments proves the existence of the psi effect. The chance probability of the overall result tends to zero as the number of experiments increases, even though the size of the effect may vary from study to study, including excursions to negative values. Accordingly, statistical psi effects are considered not reproducible but replicable. In a typical study, the mean of a stochastically fluctuating quantity is shifted by the influence of psi in the desired direction. Examples are the increase of the hit rate above its chance value when faces of playing cards are guessed (ESP) or dice are thrown (PK). Many people believe they have had experiences of this kind when they played games of chance. Statistical parapsychology provides evidence that such anomalies actually take place, even in studies conducted under controlled conditions in the lab. Without dwelling on details of mathematical analysis, Schmidt (2014) recently gave a survey of experiments on statistical psi effects. An introduction to their evaluation was published by Tressoldi and Utts (2015).

The statistically detectable psi effects appear to be everyday occurrences. They have been found at different places and by different investigators and participants. There are no regions or populations in the world that are known to be devoid of anomalous occurrences. Therefore, psi abilities are likely to be universal. However, the averages such as the hit rate of a particular study more often lie outside the confidence limits of similar studies than is allowed by chance. Obviously, the effect size scatters, which adds to the scattering of z at fixed effect size. This may seem to make it appear hopeless to search in statistical parapsychology for laws like those governing the natural sciences. Nevertheless, as in physics, measurements are taken and analyzed, and the data of the numerous investigated statistical psi effects could obey some rules. Unlike laws, these rules would allow for deviations and exceptions. Leaving aside the very rare reports on persons producing with some reliability a psi effect that far exceeds random noise, it may be worthwhile to think about a possible systematics behind the statistically detected psi effects.

The approach taken in the following is that of an applied physicist who wishes to orient himself in statistical parapsychology. In the next section, **Conjectured Rules**, I present a set of conjectured rules that possibly hold for all statistical psi effects, no matter whether it is psychokinesis (PK) or extrasensory perception (ESP). Only the data from hit-or-miss experiments with chance hit rates p = ½ and ¼ are considered, because they are particularly numerous and allow a simple and reliable analysis that can be adapted to other psi experiments whose evaluations are similar. A number of experiments on the same psi effect make up a study. The rules are based mostly on meta-analyses covering many studies of the same kind. Each single step in an experiment is a decision between hit or miss, to be called a trial in the following. An experiment consists of an arbitrary predetermined number of trials down to one. Its separation from equal experiments in a study will be of central importance in the formulation of the rules. Studies in which the separation is obviously unclear are disregarded.

Apart from the exceptions encountered in the **Calculations** section, the rules are formulated in terms of z-scores. The z-score of an experiment is defined as

$$z = \frac{k - pn}{\sqrt{p(1-p)n}}. \tag{1}$$

The nominator of the fraction is the deviation of the number k of hits from its expectation value pn, where p is the chance value of the hit rate and n the number of trials in the experiment, while the denominator is the standard deviation of k from pn. Averaging over a sufficient number N of equal psi experiments, to suppress scatter, it may be expected to result in a reasonably stable value of the actual hit rate

$$h = \sum_{j=1}^{j=N} k_j / (nN) \quad . \tag{2}$$

Using h, a mean value by definition, one may express the mean value of z by

$$< z > \; = \frac{h - p}{\sqrt{p(1-p)}} \sqrt{n} \quad . \tag{3}$$

The mean z-score is identical to Cohen's d that often, as in the present paper, serves as the definition of the effect size of an influence pushing z away from zero, its mean value for the null effect, that is in the absence of psi. The standard deviation of z from $< z >$ is taken to be that of the null effect, which is $< (z - < z >)^2 > \; = 1$. It is in general augmented by a scattering of the effect size.[1]

In the **Conjectured Rules** section, apparently common properties of the statistical psi effects are sorted out from available data and the simplest possible rules for them are formulated. Calculations associated with the rules are assembled in the four sections under **Calculations**. The first subsection, ***Unlimited Definitions of the Effect Size***, addresses the problems that may arise from the fact that the hit rate h is restricted to the interval $0 \leq h \leq 1$. Because of Equation (3), this implies that the $< z >$-score is also limited, the allowed range expanding with $\sqrt{n}$. Since the effect size per trial should in principle be unbounded, its standard definition by $< z >$ is likely to fail when $< z >$ is near one of its limits. The chance of h being so will be seen to be greatest in one-trial experiments. In order to be able to deal with such cases, a definition of the effect size is desirable that is unlimited and at small enough values of $h - p$ merges with $< z >$. Two definitions satisfying this condition and based on different models of psi effects will be proposed to convert $< z >$ into a conjectured "true" effect size and vice versa.

The two subsections ***Scattering at Small Effect Sizes per Trial*** and ***Scattering at Large Effect Sizes per Trial*** deal with the scattering of the effect size, which seems typical of psi effects. I will distinguish between weak and strong effect size scattering. It is called weak when the limits of h do not enter the calculation of shift and widening and it only widens the z-score normal distribution without affecting the shift $< z >$, while it is called strong when the limits need to be taken into account. In the case of strong scattering, the primary quantity that can be measured and calculated is the averaged hit rate, h_{av}. It is understood as the integral over a new variable h_{qu} from $-\infty$ to $+\infty$ of the product of h and its probability density, both as functions of h_{qu}. The independent variable h_{qu} will be defined by extrapolating the case of weak effect sizes. The averaged mean z-score $< z >_{av}$ is calculated from the averaged hit rate h_{av} by means of Equation (3). Based on a small dataset, the quantitative treatment of scattering is speculative. Normal distributions of the conjectured effect size will be assumed as they are common in statistics and convenient in calculations. The only histograms of effect sizes I found in the literature more or less represent a normal distribution that is wider than that of the null effect and can be explained by weak scattering. Finally, I calculate the ratios of $< z >_{av}$ in the case of strong scattering (i.e. for n = 1) to $< z >$ in the quadratic approximation (i.e. for $n >> 1$), which in most of the examples considered are less than one. In the subsection ***Comparisons with Experimental Data*** under the main Section **Calculations**, two comparisons are made between ratios of experimentally determined $< z >$-scores and the ratios of the corresponding conjectured true effect sizes, and cautious conclusions are drawn from the results. The **Conclusions** section presents a discussion of the rules and an argument as to why the small size of statistical psi effects might make sense for philosophical reasons and, if so, can probably not be substantially increased.

In a previous paper by the author, it was argued that the sizes of all psi effects are roughly equal (Helfrich 2011). However, the difference between the $< z >$-scores of PK experiments on the binary random noise generator and dream-psi, both with $p = ½$, was found to be so large (0.65 versus 0.182, see Rule 5 below) that it was tentatively attributed to the fact that the former are many-trial and the latter one-trial experiments. This guess is examined in the subsection ***Comparisons with Experimental Data*** under the main section **Calculations** of the present paper.

Conjectured Rules

For the sake of clarity, the rules are numbered. Of course, they could be differently arranged and in part decomposed or combined. They are based mainly on meta-analyses or reviews. The experimental results given here in support are in most cases not complete but a small number of examples.

Rule 1: The effect size is independent of spatial distances.

The size of psi effects does in principle not depend on the distance between participant and object or recipient and sender. There are PK studies on binary random number generators (RNGs) that show this for terrestrial distances (Dunne & Jahn 1992, 1995). The independence of distance has also been checked in ESP studies (Steinkamp 2005). In one of them a decrease of the effect size was found at large separations. No significant decrease with distance was noted in studies of remote viewing (Dunne & Jahn 2003).

Rule 2: The effect size is independent of temporal distances.

The size of psi effects does in principle not depend on differences in time between participant and object or recipient and sender. In their meta-analysis of precognition studies, Honorton and Ferrari (1989) found a dramatic decrease of the hit rate with increasing delay, which largely takes place within the first day. However, no such decay was observed in studies with selected participants. In the studies of Bem (2011), cards were guessed with the target being randomly selected only after the guess. This was interpreted as a retroactive psi effect, but PK as another possible explanation was not ruled out. In their PK studies on binary RNGs, Dunne and Jahn (1992, 1995) found no significant influence of the time shift between mental aiming and operating the RNG. The time of aiming varied from 73 hours before to 336 hours after the generation of the RNG data. The transition from PK to a kind of retroactive psi effect produced no significant break in the scatter plot of *z*-scores. An independence of temporal distances was also registered with remote viewing (Dunne & Jahn 2003).

Decreases of psi effects with increasing spatial and temporal distances are possibly due to a diminishing emotional relationship of the participant to the object or of the recipient to the sender. While the emotional relationship of a participant to an object is difficult to quantify by normal means, the decisive role of a close bond between sender and receiver was observed by Hinterberger (2008) who measured psi-induced physiological correlations at mostly large distances. Studies of the effect size as a function of distances and other parameters are much more demanding than proofs of existence of a psi effect. This is because the number of experiments has to be divided among the data points and the error of differences combines the errors of two data points.

Rule 3: The effect size is independent of the number of participants.

It does not matter much how many persons take part in an experiment, actively or passively. To employ several senders or recipients is what most physicists and engineers tend to suggest first when being told how weak psi effects are. Evidently, the lack of success of such attempts is the reason why group experiments have early on ceased to be of interest. Disappointing group studies of precognition and ESP in general were mentioned by Honorton and Ferrari (1989) and Steinkamp (2005), respectively. Dunne and Jahn (1995) found in their PK experiments that the success of pairs of participants decreased when they were equal in gender but increased when they were opposite, as compared with the success of single participants. The effect size was found to be four times larger than that of single participants when the pairs of opposite gender were "bonded," as were seven pairs in this study.

Rule 4: The effect size per experiment is independent of the predetermined number of trials in the experiments. This is on condition that on the one hand the experiments are closed, i.e. without breaks, and on the other hand clearly separated from equal or similar experiments.

The separation seems to be assured in two kinds of experiments consisting of a single trial. One of them is dream-psi (Sherwood & Roe 2003), where the temporal distance between experiments is at least a day. The other is ganzfeld-psi (Williams 2011), where the time interval between experiments is about an hour. This may not be long enough, but it seems that in addition usually the participant was changed between experiments. When an experiment consists of more than one trial, the conditions for Rule 4 seem to be well-satisfied if in a study each participant performs a single experiment that consists of a compact series of trials. Such sessions at a binary RNG that comprise roughly 20 to 50 trials are today the method of choice in many studies. In remote-viewing studies, which in addition to a recipient often involve an observer who may function as a sender, the temporal distance between experiments seems in general large enough for a clear separation, but a change of the participants would be safer. Interestingly, Baptista, Derakshani, and Tressoldi (2015) recommend that no more than one or two experiments of this kind should be carried out per day by the same participant in order to avoid fatigue or boredom. Another problem with these studies, discussed in detail by Dunne and Jahn (2003), is bringing the perception tests into the shape of yes-or-no questions with known chance probabilities.

An indirect confirmation of Rule 4 is the change in the common definition of the effect size over the course of several decades. In the beginning, it referred to the single trials, regardless of their number in an experiment, whose z-scores were thought to be independent of n. The smallness and the extreme scatter of the mean z-score thus obtained gradually led to a redefinition. Today the effect size practically always refers to whole experiments. How strictly $< z >$ is independent of n has rarely been checked. An early form of Rule 4 put forward together with experimental confirmation is the data augmentation theory (DAT) of May, Utts, and Spottiswoode (1995). According to this theory, clairvoyance unconsciously recognizes and selects rows of trials of reduced entropy in PK experiments (May et al. 1995). Numerous references to the DAT model are given in a book edited by May and Marwaha (2015).

The most convincing confirmation of Rule 4 is provided by the meta-analyses of Radin and Nelson (1989, 2000) of PK experiments with binary RNGs. They took the data from about 150 English-language references including papers published in conference proceedings, thus collecting nearly 600 experiments. In the absence of psi, the RNGs produced zeroes and ones with equal probability. The aim of the psi experiments was to mentally influence the PC so that it generates more ones than zeroes or vice versa. A most remarkable feature of these studies is the enormous range of the number of trials per experiment reaching from about 20 to 10^8. The deviation in the nominator of Equation (1) is the number of hits minus $n/2$, its chance expectation value. Without the psi effect, the scattering of the z-score results in a standard normal distribution. For this so-called null effect, the standard deviation of the z-score is $\sigma_0 = 1$, and the expectation value $E(z) = 0$, which was checked and confirmed by a histogram of 200 such experiments. With a somewhat "smoothing" assumption, which of the assignment "insignificant" made a truncated null effect distribution, Radin and Nelson in their first meta-analysis of the PK effect (1989) found a normal distribution of z-scores. Its histogram is not only shifted to $< z > = 0.65$ but also widened by a factor $\alpha = 1.5$ with respect to the null effect. In addition, there are a few outliers, while none are visible in the null-effect histogram. They were partially suppressed by a homogenization before $< z >$ and α were calculated. (However, the differences between the values of $< z >$ and α calculated after homogenization and those taken directly from the histogram seem to be small.) Knowledge of $< z >$ and α allows the computing of the entropic energies of displacement and widening, respectively. For the combination of $< z > = 0.65$ and $\alpha = 1.5$, they turn out to be practically equal (Helfrich 2011). The energies will be derived again in the subsection ***Scattering at Small Effect Sizes per Trial*** in the **Calculations** section, and their equality generalized in the subsection ***Scattering at Large Effect Sizes per Trial*** in support of Preliminary Rule 6.

A problem of PK experiments with binary RNGs is the separation of the experiments from one another: "In general, within a given reviewed report, the largest possible aggregation of non-overlapping data collected under a single intentional aim was defined as the unit of analysis (hereafter called an experiment or study)" (Radin & Nelson 1989). One would like to know if interruptions like a pause or a change of participant were excluded in these experiments. They could cause a breakup into several separate experiments. With n' being the number of effective breaks, the z-score of an experiment increases by the factor $\sqrt{n'}$ according to Stouffer's formula (see next paragraph). Therefore, breaks could be a reason for the widening.

In their second meta-analysis which in addition contained 175 new or newly found experiments, Radin and Nelson (2000) cumulated the 258 z-scores taken from PEAR (Princeton Engineering Anomalies Research) into a single one, using Stouffer's formula, $z_{\text{cum}} = \sum_{j=1}^{j=N} z_j / \sqrt{N}$. (This approximation becomes exact, apart from scattering, if $< z_j >$ can be taken to be the same for all experiments.) Was this done because decomposing the PEAR data into experiments was particularly difficult? In their second histogram of the PK effect, and in the histogram of Schub with a wider range of shown z-scores, the cumulated z-score is not marked and the indication "insignificant" is rendered simply by $z = 0$. Apart from the concentration of scores at $z = 0$ and a greatly increased roughness, there is little difference between the old and new histograms of Radin and Nelson.

The meta-analyses of Radin and Nelson (1989, 2000) were severely criticized by Bösch, Steinkamp, and Boller (2006a, 2000b) as well as by Schub (2006), who in their papers rejected its central result, an overwhelming proof of the existence of the PK effect. Radin et al. (2005a, 2005b) defended the result. In the opinion of critics, the shift of the normal distribution of hit numbers is due to a publication bias. They overlooked the fact that the widening and the outliers produced data points on both sides of the spectrum that independently of the shift drastically reduce the probability of obtaining Radin and Nelson's histogram of the psi effect by chance (Helfrich 2007).

How to be convinced that the mean shift $< z >$ is independent of the number n of trials per PK experiment at the binary RNG? First of all, the huge range of n-values leaves little room for other inferences. A partition of 377 carefully selected experiments into four practically equal blocks according to the magnitude of n provides a kind of check (Bösch, Steinkamp, & Boller 2006b). With increasing n, the authors found the decreasing $< z >$-scores 1.05, 0.75, 0.56, and 0.41. The differences, though small, may be taken to mean that the widening found by Radin and Nelson results from a superposition of normal distributions centered at different $< z >$-scores. The reason for the maximum of $< z >$ at the smallest n could be a relatively large number of interrupted experiments. This seems paradoxical, but most of the experiments with small n probably took place at a time when computer technology was nonexistent or only at its beginnings.

It should be mentioned that three PK studies with an extremely high frequency of trials (2,000,000 per 0.2 sec, once every second) produced exceptional $< z > \approx -2$, which is three times larger than what is measured at the usual 200 trials per 0.2 sec, once every second, and of the wrong sign (Ibison 1998, Dobyns et al. 2004). These results are significant but in conflict with Rule 4.

The overall effect size of the PK effect on binary RNGs obtained by Radin and Nelson (1989, 2000) in their PK meta-analyses, $< z > = 0.65$, lies on the upper border of mean z-scores of psi experiments. However, the same value was found by Honorton and Ferrari (1989) in a meta-analysis of precognition experiments. The number of experiments covered was extremely large, but in contrast to the PK studies they were quite heterogeneous. An experiment was defined as the data measured between subsequent changes of the conditions. Again, the question arises if pauses or a change of the participant occurred within an experiment because a possible breakup into shorter experiments would have caused the measured $< z >$ to be above its true value. Like Radin and Nelson, the authors found an increase of the standard deviation by a factor α as an accompanying psi effect. Before a homogenization discarding 10% of the z-scores as outliers, they obtained $< z > = 0.65$ and $\alpha = 2.48$, thereafter $< z > = 0.38$ and $\alpha = 1.45$.

Rule 5: The effect size is equal for all psi effects. Its fluctuations among studies are about as large as the average size. (However, one of the unlimited definitions of the effect size to be proposed in the following predicts a dependence of $< z >$ on the chance hit rate p according to which $< z >$ has its maximum at $p = ½$ and tends to zero for $p \to 0$ and $p \to 1$.)

All psi effects, at least those with the most common chance hit rates p = ½ and p = ¼ have similar effect sizes $< z >$. They lie preferably in or near the interval $0.2 < \langle z \rangle < 0.3$. Values below 0.1 or above 0.8 are extremely rare. Such cases call for a check if perhaps the confidence interval reaches into the preferred range. The rule applies to all modifications of ESP and PK, including retroactivity. Between studies of the same type, $\langle z \rangle$ can easily change by a factor of 2 or more, covering altogether a range whose boundaries differ by a factor of 4. In general, the limits of the confidence interval are placed at $z - \langle z \rangle = \pm 1.96\ \sigma_0$, so that the integral of the chance probability density over one or two tails of the normal distribution outside this range equals 0.025 or 0.05, respectively. In proofs of existence of a psi effect, these are the limits of significance.

Rule 5 is based on numerous meta-analyses, especially those of Radin and Nelson (1989, 2000), Honorton and Ferrari (1989), Dunne and Jahn (2003), Sherwood and Roe (2003), Williams (2011), Baptista, Derakshani, and Tressoldi (2015), Utts et al. (2010), Schmidt (2012), and Mossbridge, Tressoldi, and Utts (2012). The two last-mentioned meta-analyses deal with unconscious, physiologically detected psi effects, whose statistical evaluation was more complicated than the simple hit-or-miss scheme. We also use a comprehensive article by Bem (2011) as a source of data, even though it is not a meta-analysis. It describes nine studies, each with usually about 100 participants, of various retroactive psi effects, that were guessing tasks with the targets being randomly selected after the guessing. A recent meta-analysis by Bem et al. (2015) covers these studies and, as a check, 81 similar ones. According to the authors, the effect sizes of the additional studies as a whole, are smaller than or practically equal to Bem's values, depending on the method of analysis. I do not further discuss them because not all of them are based on hit-or-miss trials and effect sizes are expressed in a different measure (Hedge's *g*).

Moreover, it appears appropriate to include the *z*-score of Nelson's Global Consciousness Project (Nelson 2001, Nelson et al. 2002). The *z*-scores z_{gcp} are expressed in terms of a sum of the type

$$\sum_{i=1}^{n} (z_i^2 - 1) / \sqrt{\langle (z^2 - 1)^2 \rangle M} = \sum_{i=1}^{M} (z_i^2 - 1) / \sqrt{2M},$$

where M is a very large number that increases with the time elapsed since the start of the experiment. On the basis of 500 experiments, Nelson reports $\langle z_{\text{gcp}} \rangle = 0.3269$, which is in the typical range of effect sizes (as of December 2015, see GCP updates online, noosphere.princeton.edu). In these equations, z_i and z designate measured and chance values, respectively, the average being taken over the latter.

In their meta-analysis of dream-psi, Sherwood and Roe (2003) distinguish two periods. The first comprises the experiments carried out at

Maimonides Medical Center Brooklyn from 1962 to 1978. It is characterized by many exploratory experiments and the preferred use of telepathy as ESP channel. The hit rate of the 450 experiments was 63% instead of a 50% chance probability. The second period, called post-Maimonides, lasted from 1977 to 2007. The 820 additional experiments differed in location and method, and their effect size was smaller than that of the first period. In some cases, a large number of recipients simultaneously received the same dream content by telepathy from the same sender. The number of experiments was equated to that of the number of recipients. However, according to Rule 4, it should rather be one because there was a single sender. For this reason, we prefer $< z > = 0.26$, the value of the first period, over $< z > = 0.182$, the value calculated by Radin (2006) for the total of 1,270 experiments.

The $< z >$-scores of Bem's studies (2011) varied within the range given above. In most studies, a test consisting of two questions distinguished between stimulus-seeking and other participants. The $< z >$-scores were computed for both groups and for the total of participants in a study. The stimulus-seekers were clearly more successful than the others, the averaged effect sizes of the groups being $< z > = 0.43$ and 0.10, respectively. The overall effect size was $< z > = 0.22$.

How to optimize psi effects with respect to size and replicability is the main subject of a meta-analysis of Baptista, Derakshani, and Tressoldi (2015). They consider ganzfeld-psi, card guessing, remote viewing, and dream-psi. The most important precondition for large effect sizes appears to be selection of the participants. Belief in the existence of psi effects, experience with psi experiments, success in previous such experiments, and training in meditation all are helpful. The $< z >$-scores of Bem's studies show that being a stimulus-seeking person can be enough to achieve above-average effect sizes. The aforementioned small but significant decrease of $< z >$ with increasing n, as noted by Bösch, Steinkamp, and Boller (2006b) in the data of Radin and Nelson, could be explained not only by breaks but alternatively (and less likely) on the basis of Rule 5 by a predominance of enthusiasm in the shorter, early PK experiments and a predominance of routine in the longer, later ones.

Preliminary Rule 6: The *z*-scores of very large numbers of experiments carried out by different groups and over a long period of time tend to end up in normal distributions.

The size of psi effects is not constant but undergoes fluctuations from study to study. Ganzfeld-psi represents a well-investigated example of the type $n = 1$, as demonstrated, e.g., by the meta-analyses by Williams

(2011) and Baptista, Derakshani, and Tressoldi (2015). The $< z >$-scores of these studies more often are outside the limits of confidence of similar studies than would be expected on the basis of null-effect scattering. The easiest way of recognizing fluctuations of the effect size is to look at the experimental standard deviation σ of z which in their presence exceeds that of the null effect, i.e. $\sigma > \sigma_0 = 1$. If there is a widening of the standard deviation, additional effort is required to gain information on the effect-size distribution causing it. Although indications of a scattering of the effect size have often been observed, there seem to be no systematic investigations of the affected z-score distribution functions, apart from Radin and Nelson's (1989, 2000) meta-analyses and Schub's critique thereof.

In the subsections ***Scattering at Large Effect Sizes per Trial*** and ***Comparisons with Experimental Data*** in the **Calculations** section, I will presuppose normal distributions of the scattered effect size, thereby permitting a lowering of the effect size by homogenization and elimination of outliers. Calculations with an acceptable effort are possible only with normal distributions. From the experimental point of view, the assumption that they are at least reasonable approximations can be inferred only from Radin and Nelson's meta-analyses, the criticisms of which have been pointed out above. The same applies to the assumption that the energies of shifting and widening the distribution of the z-scores are equal or proportional to one another.

Calculations

Unlimited Definitions of Effect Size

The starting point of all calculations is the binomial distribution. I consider n equal trials of the same chance hit rate p. It does not matter at this point whether they belong to a single experiment with n trials or a series of n equal one-trial experiments. The possible total numbers of hits are $k = 0$, 1 2, ... , n. The probability of exactly k hits may be expressed by the term $B_{n,p}(k)$ of a binomial distribution:

$$B_{n,p}(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \tag{4}$$

The sum of these terms over k satisfies for all n

$$\Sigma_{k=0}^{n} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = 1. \tag{5}$$

It is, as may be said, normalized to unity.

The expectation value of the hit number is np. The term $k = np$ (or the one next to np) is the term with the largest probability. According to the DeMoivre-Laplace theorem, the binomial distribution asymptotically approaches, for $n \to \infty$ but some fixed x in $k = np + x\sqrt{np(1-p)}$ a normal distribution of the probability density

$$w_{n,p}(k) = \frac{1}{\sqrt{2\pi p(1-p)n}} \exp\left[-\frac{(k-np)^2}{2p(1-p)n}\right] . \tag{6}$$

Its integral over k is equal to unity. This holds exactly only if the integral reaches from $-\infty$ to $+\infty$. In the present case it is restricted to the interval $0 \leq k \leq n$. Equation (6) can also be read as a discrete probability function, the sum over all integers k tending to unity for $n \to \infty$. Insertion of Equation (1) into Equation (6) leads to the standard form of the normal distribution

$$w(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2 / 2), \tag{7}$$

where z is usually regarded as a continuous variable.

The chance probability of obtaining, at the chance hit rate p with n trials nh hits, is given by $B_{n,p}(nh)$. The quantity of interest is the ratio $Q_{n,p}(h)$ of the probabilities of this state to that of the ground state, i.e. the most probable state. All states $B_{n,p'}(nh')$ with arbitrary $0 < h' < 1$ lend themselves as ground states. The probabilities of these ground states are not exactly equal but differ by the factor $1/\sqrt{h'(1-h'}$. The natural choice may seem to be $h' = p$ so that ground state and excited state, i.e. state of lower probability, have the same chance hit rate p. However, the more convenient choice is $h' = h$ because then all the factorials cancel each other. This leads immediately to

$$\frac{B_{n,p}(nh)}{B_{n,h}(nh)} = \frac{p^{nh}(1-p)^{n(1-h)}}{h^{nh}(1-h)^{n(1-h)}} . \tag{8}$$

It does not matter that the binomial terms $B_{n,p}(nh) B_{n,h}(nh)$ belong to different binomial series, because the sum of each series is normalized to unity. Taking logarithms, one may write

$$\ln \frac{B_{n,p}(nh)}{B_{n,h}(nh)} = -n\eta(p,h), \tag{9}$$

where

$$\eta(p,h) = (1-h)\ln\frac{1-h}{1-p} + h\ln\frac{h}{p}. \tag{10}$$

This function of h and p has its minimum at $h = p$ with $\eta(p,p) = 0$. If one chooses $B_{n,p}\,(np)$ instead of $B_{n,h}\,(nh)$ as the ground state, one has to add on the right side of Equation (9) the term $(1/2)\ \ln\{p(1-p)\ /\ [h\ (1-h)]\}$. This follows from Stirling's formula applied to $B_{n,h}\,(nh)\ /\ B_{n,p}\,(np)$ as well as from the factors preceding the exponentials of the probability normal distributions. The additional term does not depend on n, thus being negligible at large n. It disappears if instead of the probability ratios between the individual states of maximum probability one considers the ratios of the products of this probability and the respective standard deviation (or a fraction thereof). Moreover, it is completely avoided when Equation (10) is deduced on the basis of a single-trial approach, as will be done below (see Equation (18)). Equation (10) without the second term can also be derived in terms of the physics of the isothermal ideal gas[2]. In a slightly modified form, it is a special case of the Kullback-Leibler distance or relative entropy. Accordingly, the formula of the probability ratio to be employed in the following is

$$Q_{n,p}(h) = \exp[-n\boldsymbol{\eta}(p,h)\} \tag{11}$$

In statistical thermodynamics, the probability of a state being occupied is proportional to the Boltzmann factor, another exponential function. Its exponent is $-E/k_BT$, where E is the energy of the state, k_B Boltzmann's constant, and T the absolute temperature. Obviously, $-\ln\eta(p,h)$ may be interpreted as the entropic free energy of a state minus that of the ground state, both per trial and divided by k_BT. A temperature dependence of psi effects is not known. If E/k_BT is independent of temperature, E must be proportional to k_BT. For convenience, $\eta(p,h)$ will sometimes simply be called energy.

The function $\eta(p,h)$ represents the energy per trial required to bring the hit rate from the chance value p to the actual value h. It can be expanded into a power series of $(h-p)$. Omitting the terms of higher than quadratic order in $(h-p)$, one obtains

$$\eta_{\text{qu}}(p,h) = \frac{(h-p)^2}{2p(1-p)}, \tag{12}$$

i.e. the quadratic approximation of $\eta(p,h)$. Comparison of Equation (12) with Equation (1) leads to

$$\eta(p,h) = \frac{1}{2} < z_{n=1} >^2 , \tag{13}$$

the subscript $n = 1$ indicating one-trial experiments. The function $\eta(p,h)$ is defined only in the interval $0 \le h \le 1$ where it is finite everywhere. However, the derivatives of $\eta(p,h)$ with respect to h diverge at the limits of h. The first two derivatives are

$$d\eta(p,h) / dh = \ln[h(1-p) / p(1-h)] , \tag{14}$$

and

$$d^2\eta(p,h) / dh^2 = 1 / h(1-h). \tag{15}$$

Also of interest will be the first derivative of $\eta_{\text{qu}}(p,h)$

$$\begin{aligned} d\eta_{\text{qu}}(p,h) / dh &= (h-p) / p(1-p) \\ &= < z_{\text{n=1}} > / \sqrt{p(1-p)} . \end{aligned} \tag{16}$$

For small enough $|h-p|$, the functions $\eta_{\text{qu}}(p,h)$ and $\eta(p,h)$ are practically identical. The definition of the effect size by $<z_{\text{n=1}}>$ becomes questionable to the extent that $\eta(p,h)$ and its quadratic approximation $\eta_{\text{qu}}(p,h)$ differ from each other. Any redefined effect size should be unlimited but merge with the quadratic approximation at small sizes. Two modified definitions of the effect size satisfying these requirements are proposed next. They are based on physically inspired concepts of the psi effects that might be called field model and momentum model.

Beginning with the field model, let me imagine the psi effect to be caused by the psi field

$$\kappa(p,h) = \frac{d\eta(p,h)}{dh}, \tag{17}$$

In equilibrium, a psi field $\kappa > 0$ shifts the minimum of the total one-trial energy $\eta(p,h) - \kappa(p,h)h$ from $h = p$ to some $h > p$. The quantity $\kappa(p,h)$ is reminiscent of a physical force. However, a force can be defined as the negative derivative of energy with respect to length, while $\kappa(p,h)$ is the positive derivative with respect to hit rate. The dimension of $\kappa(p,h)$ is again energy in units of k_BT as h is a dimensionless quantity. Its meaning becomes apparent by expanding the fraction in Equation (17) by a sufficient number n of trials and writing the result as an equation of differences, $\Delta(n\eta(p,h)) = \kappa(p,h)\,\Delta(nh)$. The number of hits must be a natural number between 0 and n. A decrease of this number by 1, i.e. $\Delta(nh) = 1$, is accompanied by the release of the energy $\kappa(p,h)$ which at equilibrium is exactly what is absorbed by the system when a miss is converted into a hit. The fact that κ is a released energy implies that the ratio of the probability of a trial being a hit to that of being a miss or, in other words, the rate of hits to the rate of misses, is $pe^{\kappa}/(1-p)$, where $e^{\kappa} = \exp[-(-\kappa)]$ represents the "Boltzmann factor" of the energy, $-\kappa(p,h)$. This results in the following formula for h:

$$h(p,\kappa) = pe^{\kappa} / [pe^{\kappa} + (1-p)] = 1/(1+\frac{1-p}{p}e^{-\kappa} \quad . \qquad (18)$$

Solving Equation (18) for $\kappa(p,h)$ that is subsequently substituted by Equation (17), leads, in fact, back to Equation (14) and finally Equation (10). The last form of Equation (18) serves to show that it is easier to compute h as a function of κ than the other way round.

The limited hit rate h is a one-to-one function of unlimited $\kappa(p,h)$ or $\sqrt{p(1-p)}\kappa(p,h)$ that merges with its quadratic approximation (see Figure 1 below). In the quadratic approximation, the psi field $\kappa(p,h)$ takes the form

$$\kappa(p,h) = \frac{d\eta(p,h)}{dh} = \frac{h-p}{p(1-p)} = \; < z_{\mathrm{n=1}} > /\sqrt{p(1-p)} \qquad (19)$$

that is bounded because of the limits of h. In an x/y plot, the psi field shifts the parabolic potential $\eta_{\mathrm{qu}}(p,h)$ over the (horizontal) distance $(h-p)$, thereby lowering without deforming it.

The effect size can be expressed by $\sqrt{p(1-p)}\kappa(p,h)$ or by the field $\kappa(p,h)$. Although direct use of the psi field κ as effect size may seem attractive (and is made in Figure 1), practical reasons argue for keeping $\sqrt{p(1-p)}\kappa(p,h)$, as long as it is not known which of the two variants, if any, is the correct one. Otherwise, all effect sizes $< \mathrm{z} >$ reported in the literature would have to be magnified by the factor $1/\sqrt{p(1-p)}$ to convert them into κ.

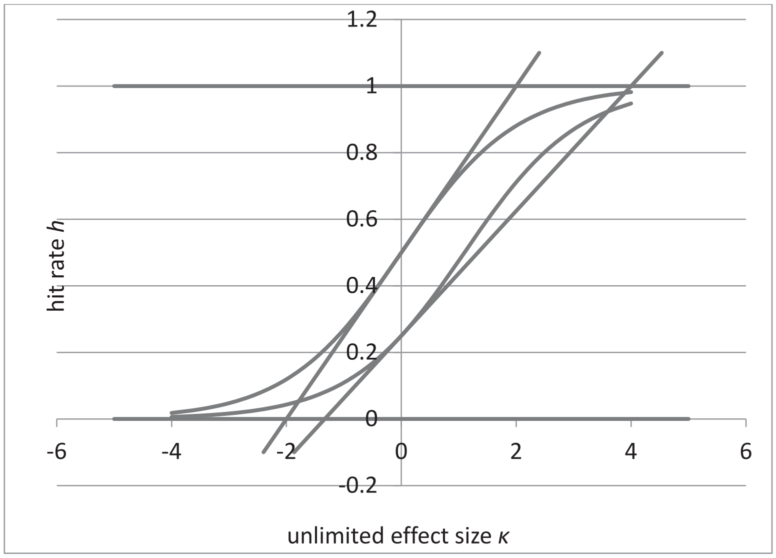

Figure 1. Hit rate *h* as a function of the psi field $\kappa = d\eta(p,h)/dh$. The curved lines represent the dependence of h on κ when the exact energy $\eta(p,h)$ is used. In the approximate calculations, each curved line is replaced by three pieces of straight lines. The central one of them derives from the quadratic approximation $\eta_{qu}(p,h)$ of the exact energy, the horizontal ones represent the limiting hit rates 0 and 1. The two structures refer to $p = ½$ (left) and $p = ¼$ (right). The psi field κ equals $(h_{qu} - p) / p(1 - p)$ (see main text).

A more comfortable alternative would be to use $\kappa/2$ as the new effect size. The correction factor would then be $1/[2\sqrt{p(1-p)}]$, so that the numbers do not change for $p = ½$, the most often investigated case. Whenever $p \neq ½$, both κ or $\kappa/2$ are larger by this factor than their values at $p = ½$. For $p = ¼$, the factor is 1.15.

In the other model allowing effect sizes of unlimited size, the psi effect is caused by a momentum s that is the new effect size. It could be carried by a particle with the kinetic energy $s^2/2m$ hitting the system at a particular trial. The mass m is equated to unity so that the maximum transferable energy is $½\, s^2$. At the beginning, the system is thought to be in the ground state with the potential $\eta(p,p) = 0$. The momentum excites the z-score of the trial in the direction of its sign and is assumed to be completely absorbed by the system if the hit rate associated with the energy does not exceed the limits $h = 1$ or $h = 0$. Within this range, s is defined by the equation

$$s = \pm\sqrt{2\eta(p,h)} \quad . \tag{20}$$

If the quadratic approximation holds, the momentum obeys

$$\frac{1}{2}s^2 = \eta\ (p,h) = \frac{(h-p)^2}{2p(1-p)} = \frac{1}{2} < z_{n=1} >^2 \tag{21}$$

Beyond its mergers with one of the limits of h, $h(p,s)$ is assumed to continue on one of the straight lines representing h = 1 or 0, depending on whether s increases or decreases, respectively (see Figure 2 below). This implies trimming away any parts of $h(p,s)$ that exceed the limits of h. A questionable simplification of the momentum model is the disregard of the null-effect fluctuations. Despite the apparent superiority of the field model, I will continue to consider both models, because it is an entirely open question how psi works. Neither model removes the mystery from psi, they only move it to an earlier moment in the chain of events.

When could it be necessary to go beyond the quadratic approximation? Probably only when the hit rate h associated with κ or s is near its limits at 1 and 0. From now on, I will distinguish between the number of trials, n, and the number of experiments, N. The condition just stipulated is certainly not satisfied by experiments consisting of many trials. According to Rule 4, the effect size per experiment, $< z >$, may be expected to break up into an effect size per trial of $< z > / \sqrt{n}$, a quantity that rapidly decreases with the number of trials. A rise of n from 1 to 2 already makes a great difference. Therefore, the limits of h interfere the most in one-trial experiments, the only case of low n to be considered in the following. Incidentally, fields and momenta varying with $1/\sqrt{n}$ result automatically if the total values are decomposed into n equal components in n-dimensional space. This might be interpreted as a physical explanation of Rule 4.

The dependences of the hit rate on the effect sizes κ and s in the case $n = 1$ are shown in the two figures for the two most common chance hit rates, $p = ½$ and $p = ¼$. Figure 1 depicts the functions $h(p,\kappa)$, their curvilinear plots approaching the limits of h without ever reaching them. Figure 2 depicts the functions $h(p,s)$. Their plots are curvilinear as long as the energy $\eta(p,s)$ associated with s can be fully absorbed by the system. They are assumed to change to the horizontal straight lines representing the limits of h at 1 and 0 where they merge with them rather abruptly. This is the basic version of the momentum model; two more complicated variants will be briefly considered in the subsection ***Scattering at Large Effect Sizes per Trial***. For small $|h-p|$

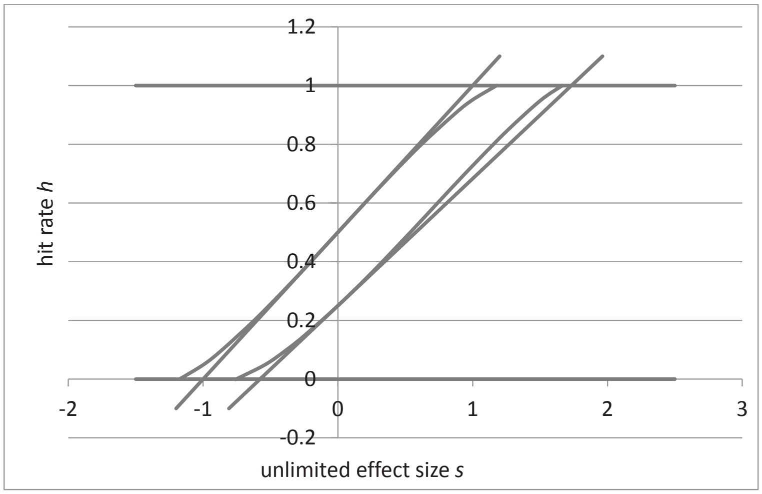

Figure 2. Hit rate *h* as a function of the psi momentum *s*, obtained by plotting $s(p,h)=\sqrt{2\eta(p,h)}$ as a function of *h*. The curved lines represent the dependence of *h* on *s* when the exact energy *η*(*p,h*) is used. Their mergers with the horizontal lines *h* = 1 and *h* = −1 are too abrupt to be resolved in the Figure. The straight-lined approximations of *h*(*p,s*) correspond to those in Figure 1. The two structures refer to *p* = ½ (left) and *p* = ¼ (right). The psi momentum *s* equals $(h_{\text{qu}}-p)/\sqrt{p(1-p)}$ (see main text).

the curvilinear functions are seen to merge with the corresponding linear dependences of the quadratic approximation. According to Figures 1 and 2, it is only near the limits of *h* that the new effect sizes deviate appreciably from $< z_{\text{n=1}} >$, the measured mean *z*-score. For $p = ½$, the deviation in the field model is circa +15% at $(h-p) = 0.3$ and circa +40% at $(h-p) = 0.4$, in the momentum model it hardly exceeds +10%. For $p = ¼$ in the field model it may be negative, reaching circa −20% at $(h-p) = 0.5$, but from there it rises to the positive side of $< z_{\text{n=1}} >$. The positive or negative deviations in the momentum model are in general smaller than those in the field model. However, beyond the mergers the hit rate *h*(*p,s*) does not respond to further increases of the effect size, while in the field model the limited hit rate is a one-to-one function of the unlimited effect size.

The effect sizes of most one-trial experiments reported in the literature are so small that to a good approximation they can be expressed by $< z_{\text{n=1}} >$ as measured. However, a substantial downward or upward deviation of the conjectured effect size from $< z_{\text{n=1}} >$ is still possible if the effect size scatters so widely that part of its spectrum lies outside the range of validity of the quadratic approximation. Dealing with the scattering of the effect size is the next and final task. Two cases will be distinguished: weak scattering that can be treated within the quadratic approximation and strong scattering that cannot.

The scattering of the effect size is probably composed of three parts. In the first place, the psi-ability or psi-sensitivity has been found to vary considerably among participants. Also, personal sensitivity can change over the course of time and with the circumstances. The investigators and checkers may exert another influence. In addition, there may be fluctuations of the effect size caused by external influences that do not depend on the persons involved and may be inexplicable. Technical irregularities can arise from errors in the counting of the experiments. A clear distinction of these sources is not possible. Therefore, the total scattering will be represented by normal distributions in the following calculations.

Scattering at Small Effect Sizes per Trial

The meta-analyses of Radin and Nelson (1989, 2000) start from the standard normal distribution of z-scores for the null effect at large n, as described by Equation (7). With a simple mathematical ansatz, one can reproduce the unintended widening of this distribution by a factor α that emerges in the meta-analysis of the PK effect in addition to the intended shift. It is sufficient to assume that the effect size scatters and that the scattering obeys a normal distribution (Helfrich 2011)

$$w(\zeta) = \frac{1}{\tau\sqrt{2\pi}} \exp(-\zeta^2 / 2\tau^2), \tag{22}$$

where ζ is the variable part of the effect size expressed in units of z. Combining Equation (7), i.e. the standard normal distribution associated with the scattering of the null effect, and Equation (22), one obtains the normal distribution

$$\begin{aligned} w(z) &= \int_{-\infty}^{\infty} \frac{1}{2\pi\tau} \exp[-\frac{(z - < z > -\zeta^2)}{2} - \frac{\zeta^2}{2\tau^2}] d\zeta \\ &= \frac{1}{\sqrt{2\pi(1+\tau^2)}} \exp[-\frac{(z - < z >)^2}{2(1+\tau^2)}], \end{aligned} \tag{23}$$

The standard deviation of effect-size scattering, τ, is still unknown, but comparison of Equation (23) with the result of Nelson and Radin's meta-analysis immediately leads to

$$\alpha^2 = 1 + \tau^2. \tag{24}$$

Insertion of α = 1.5 yields τ = 1.18. Evidently, the normally distributed effect size scattering does not affect the measured values of $< z >$ and h, despite the fact that they are averages of mean values in its presence.

How large is the probability ratio $\varphi(< z >,\alpha)$ per experiment for the transition from distribution Equation (7) to distribution Equation (23)? The probability ratio of a transition from $z = 0$ to a particular z-score in the widened distribution is $\alpha \exp(-(z^2 - 1) / 2)$. The factor $\alpha > 1$ takes into account that more states, i.e. k-values, are available per standard deviation in the new than in the old distribution, provided the z-scale is retained. The z^2 term is averaged over the new distribution and $< z^2 >$ substituted by means of the well-known relationship $\alpha^2 = < z^2 > - < z >^2$. The result is

$$\varphi(< z >, \alpha) = \exp[\frac{\ln \alpha^2 - (\alpha^2 - 1) - < z >^2}{2}]. \tag{25}$$

The number 1 in the nominator of the exponent ensures that $\varphi(< z >,\alpha) = 1$ if new and old distributions are identical. Evidently, the last term in the exponent is the energy of shifting the normal distribution, while the preceding terms represent the energy of its widening. Both are, of course, energies in units of $k_B T$. They happen to be practically equal for the combination of $< z >$ = 0.65 and α = 1.5, the values of PK on binary RNGs taken from Radin and Nelson's (1989) meta-analysis. Insertion into Equation (25) yields $< z >^2 / 2$ = 0.22 and $[\ln\alpha^2 - (\alpha^2 - 1)] / 2 = 0.21$. Exact equality means

$$(a^2 - 1) - \ln \alpha^2 = < z >^2 . \tag{26}$$

The derivation of the left side of this equation with respect to the widening $(\alpha - 1)$ at the point where both sides of Equation (26) are zero leads to the linear relationship

$$(\alpha - 1) = < z > / \sqrt{2}, \tag{27}$$

Over a surprisingly wide range of $< z >$, it is a good approximation to Equation (26), the deviation of $(\alpha - 1)$ as calculated from Equation (27) relative to the value obtained from Equation (26) reaching hardly +5% at $< z > = 1$. One may wonder whether widening and shift, i.e. $(\alpha - 1)$ and $< z >$, are equal or proportional to each other rather than the associated energies.

Another question is whether τ should perhaps be 1 instead of 1.18, so that the standard deviations of null-effect and effect-size scattering are equal. This would be exactly valid, e.g., for the combination of $< z > = 0.65$ and $\alpha = 1.46$. The uncertainties of the reported experimental data, including those related to homogenization, allow for many hypotheses. Additional problems may arise if according to some external criterion the participants in a study or meta-analysis can be divided into groups with different $< z >$-scores. In dealing with the scattering of large effect sizes, I will ignore all of these possibilities and adhere to the assumption that the energies of shifting and widening the normal distribution of the null effect are equal.

Scattering at Large Effect Sizes per Trial

The aim of these especially speculative and approximate calculations is to predict the averaged hit rate h_{av} and the averaged mean z-score $< z >_{av}$ computed from h_{av} by means of Equation (3). The subscript av serves to distinguish these theoretical numbers from measured hit rates and effect sizes. For this purpose, an idea of the statistics governing the effect sizes κ and s is needed: Let me simply assume the quadratic approximation to continue beyond the limits of h, calling the new, unlimited variable h_{qu}. The approach seems plausible because externally controlled psi fields or momenta should not depend on the properties of the system on which they act. For the reasons given above, only one-trial experiments will be considered. The associated energy is

$$\eta_{qu}(p,h_{qu}) = \frac{(h_{qu}-p)^2}{2p(1-p)} \quad . \tag{28}$$

The independent variable h_{qu} practically coincides with h as long as the quadratic approximation is applicable. According to Equation (28), the conjectured psi field is

$$= \frac{d \quad _{qu}(p,h_{qu})}{dh_{qu}} = \frac{h_{qu} \quad p}{p(1 \quad p)} \tag{29}$$

while the conjectured psi momentum is

$$s = \sqrt{2\eta_{qu}(p,h_{qu})} = \frac{h_{qu}-p}{\sqrt{p(1-p)}} . \tag{30}$$

Note the linear relationships of h_{qu} with κ and s.The new variable, and thus κ and s, are thought to be normally distributed.The central value h_0 of the h_{qu} distribution can be determined experimentally, if it is possible to do equal or similar many-trial experiments with n so large that the quadratic approximation holds. According to Rule 4 the effect sizes of one-trial and many-trial experiments should be equal, which implies $h_0 - \mathrm{p} = < z_{n\gg 1} > \sqrt{p(1-p)}$. In analogy to Equation (23), the probability density of h_{qu} is expressed by

$$w_{\mathrm{p}}(h_{\mathrm{qu}}) = \frac{1}{\tau\sqrt{2\pi p(1-p)}} \exp[-\frac{(h_{\mathrm{qu}} - h_0)^2}{2p(1-p)\tau^2}]. \qquad (31)$$

The averaged hit rate h_{av} is the integral over h_{qu} from $-\infty$ to $+\infty$ of the product of this function and $h(p,\kappa)$ or $h(p,s)$. While $h(p,\kappa)$ is rendered in explicit form by Equation (18), there is no explicit form of the function $h(p,s)$. The actual hit rate as a function of h_{qu} will be designated $h_{\mathrm{p}}(h_{qu})$, the model to which it applies following from the context.

In view of the speculative character of the models and the uncertainties of the experimental data, it seems reasonable in a first, approximate calculation of the averaged hit rate to prefer transparency over mathematical rigor. Therefore, $h(p,\kappa)$ and $h(p,s)$ are replaced by the three straight sections representing the straight-lined approximations of the models (see Figure 1 and Figure 2). Coming from $h_{qu} = -\infty$ on the straight line $h = 0$, one changes at the intersections to the straight lines of the quadratic approximation and from there to the line $h = 1$ on which one continues up to $h_{qu} = +\infty$. The resulting function is $h = h_{qu}$ in the interval $0 \le h_{qu} \le 1$, while it is 0 for $h_{qu} \le 0$ and 1 for $h_{qu} \ge 1$. The functions $h(p,s)$ of the basic version of the momentum model deviate only in their curvilinear parts from this approximation, while the functions $h(p,\kappa)$ differ everywhere.

In the case of the momentum model, one has to make assumptions on how to deal with the momenta that cannot be fully absorbed by the system. Three simple choices are to be considered: Excessive values of h_{qu} are either trimmed off to the next absorbable value, thus becoming $h = 0$ or 1, as was done above to define the basic version of this model, or they are lost and the loss is compensated so that the integral of the probability density remains unity. Trimming is the simplest method and more compatible with physics than the other two. Compensation is achieved by assuming for the lost parts of the spectrum the value of the null effect, $h = p$, or by multiplying what is left of the spectrum by a renormalization factor. The two variants of the momentum model with lost but compensated excessive momenta are included because they yield the strong reduction of $< z >$ in single-trial experiments that was originally inferred from the experimental data. However, they require modifications of the function $h(p,s)$ representing the basic, i.e. trimmed, version of this model. If the null effect serves as compensation, h cannot remain at the values 1 or 0 once these limits are attained. Instead, both values have to be substituted by p beyond the mergers or, in the rectilineaer approximation, intersections with the lines $h = 1$ or 0. Such a breakdown of the psi effect beyond its extrema would mean that excessive momenta pass the system without leaving a trace. While this appears unlikely, it cannot be entirely ruled out on the basis of presently available data. The renormalization variant requires modifications even less acceptable from the physics point of view.

Three integrals of the probability density, Equation (31), which can be regarded as areas, are needed for the calculation of h_{av}, the averaged hit rate. The areas between $h_{\mathrm{qu}} = -\infty$ and $h_{\mathrm{qu}} = 0$, $h_{\mathrm{qu}} = 0$ and $h_{\mathrm{qu}} = 1$, $h_{\mathrm{qu}} = 1$ and $h_{\mathrm{qu}} = +\infty$ will be called A_L, A_M, A_R, respectively. For instance,

$$A_R = \int_1^{\infty} \frac{1}{\tau\sqrt{2\pi(1-p)p}} \exp[-\frac{(h_{\mathrm{qu}}-h_0)^2}{2(1-p)p\tau^2}]dh_{\mathrm{qu}}. \qquad (32)$$

Of course, the sum of the three areas equals unity. The integrals A_L and A_R are the probabilities of the hit rates $h = 0$ and $h = 1$, respectively. The integral of the product of h_{qu} and the probability density, Equation (31), from $h_{\mathrm{qu}} = 0$ to $h_{\mathrm{qu}} = 1$,

$$h_M = \int_0^1 h_{\mathrm{qu}} \frac{1}{\tau\sqrt{2\pi(1-p)p}} \exp[-\frac{(h_{\mathrm{qu}}-h_0)^2}{2(1-p)p\ \tau^2}]dh_{\mathrm{qu}}, \qquad (33)$$

is the (unrenormalized) contribution to h_{av} of the interval between $h = 0$ and $h = 1$. An elementary integration yields

$$h_M = -\sqrt{\frac{p(1-p)\tau^2}{2\pi}} \left\{ \exp[-\frac{(1-h_0)^2}{2p(1-p)\tau^2}] - \exp[-\frac{h_0^2}{2p(1-p)\tau^2}] \right\} + h_0 A_M. \qquad (34)$$

Combining the contributions, one obtains in the field model and the momentum model with trimming the averaged hit rate

$$h_{\mathrm{av}} = h_M + A_R, \qquad (35)$$

where the second term represents the contribution of the straight line $h = 1$. In the variants of the momentum model where excessive momenta are lost, one finds

$$h_{\mathrm{av}} = h_M + p(1 - A_M), \tag{36}$$

if the loss is compensated by the null effect, and

$$h_{\mathrm{av}} = h_M \,/\, A_M \;, \tag{37}$$

if it is compensated by renormalization.

Two sets of averaged hit rates calculated from Equations (35) to (37) and ratios $\mathrm{R} = \dfrac{h_{\mathrm{av}} - p}{h_0 - p} = \dfrac{< z_{\mathrm{n=1}} >_{\mathrm{av}}}{< z_{\mathrm{n>>1}} >}$ where $< z_{\mathrm{n=1}} >_{\mathrm{av}} = (h_{\mathrm{av}} - p) \sqrt{p(1-p)}$ is the averaged $< z >$-score, are listed in Table 1. The ratio R serves as the correction factor of $< z_{\mathrm{n}\gg 1} >$ that brings it down (or up) to $< z_{\mathrm{n=1}} >_{\mathrm{av}}$, it is unity in the quadratic approximation. The calculations refer to the value pairs $< z > = 0.65$ with $\alpha = 1.5$ measured at $n » 1$ and $< z > = 0.4066$ with $\alpha = 1.3$. The first pair nearly and the second one exactly satisfy Equation (26). Only positive $< z >$-scores at the chance hit rates $p = ½$ and ¼ are considered. For $p = ¼$, the corrections are little, i.e. R remains close to unity. There are small reductions (R < 1) for $< z > = 0.65$ and small enhancements (R > 1) for $< z > = 0.4066$. The latter are due to the fact that for $h_0 < 0.5$, a larger part of the scattering spectrum lies on the left of the range $0 < h_{\mathrm{qu}} < 1$ than on the right. More interesting are the results for $p = ½$, where R is generally reduced, apparently tending to unity with decreasing $(h_0 - p)$. The reductions listed in Table 1 are particularly distinct in the case $< z > = 0.65$, ranging from 0.60 in the approximate field model and the momentum model with trimming to R = 0.13 in the momentum model with compensation by the null effect.

In the field model, the hit rate $h_{\mathrm{p}}(h_{\mathrm{qu}})$ deviates markedly from the straight-lined approximation. Therefore, exact integrations over h_{qu} of the products $w_{\mathrm{p}}(h_{\mathrm{qu}})h_{\mathrm{p}}(h_{\mathrm{qu}})$ were done in addition (by online integration). The values of h_{av} and R thus calculated are in the last column of Table 1, next to those obtained for the field model and the trimmed momentum model in the straight-lined approximations. For $p = ½$ the averaged hit rates in the last column are smaller by about 10% than those of the approximation. For $p = ¼$ they are larger by up to 20% in the range between $h_{\mathrm{qu}} = ¼$ and $h_{\mathrm{qu}} = 1$, probably because of the pronounced upward bulge of $h_{\mathrm{p}}(h_{\mathrm{qu}})$ relative to $h = h_{\mathrm{qu}}$ that is visible in Figure 1.

TABLE 1

Calculated Approximate Values of h_{av} and R = $< z > / < z_{n>>1} >$ = $(h_{av} - p) / (h_0 - p)$ for the Three Variants of the Momentum Model and the Field Model

		Momentum model with approximate rectilinear $h(h_{qu}) = 0$, h_{qu} or 1 — Excessive effect sizes: lost but compensated by null effect	lost but compensated by renormalizing	trimmed	Field model* with exact curvilinear $h(h_{qu})$
$p = 1/2$	h_{av}	0.542	0.576	0.696	0.677
$< z > = 0.65$, $\alpha = 1.5$	R	0.13	0.23	0.60	0.54
$p = 1/2$	h_{av}	0.557	0.580	0.653	0.634
$< z > = 0.4066$, $\alpha = 1.3$	R	0.28	0.39	0.75	0.66
$p = 1/4$	h_{av}	0.431	0.510	0.522	0.551
$< z > = 0.65$, $\alpha = 1.5$	R	0.64	0.92	0.97	1.07
$p = 1/4$	h_{av}	0.427	0.464	0.439	0.475
$< z > = 0.4066$, $\alpha = 1.3$	R	1.00	1.21	1.07	1.28

* The values of the momentum model with trimming in the next-to-last column apply, as approximations, also to the field model.

One would like to know which of the proposed models, if any, is correct or at least makes the best predictions. From the physicist's point of view, the field model appears much more attractive than any of the three variants of the momentum model. The variant with trimming is the most plausible choice among them. However, the excessive parts of the conjectured true effect size are trimmed away and not detectable. The following comparisons of theory and experiment are an attempt to find the most likely model.

Comparisons with Experimental Data

As already mentioned, in an earlier paper by the author the question arose whether the limits of the hit rate at 0 and 1 can diminish the measured $< z >$ in comparison to what it would be without this limitation (Helfrich 2011). The reason was the great difference between $< z > = 0.65$ as obtained in the meta-analyses of PK experiments on binary RNGs by Radin and Nelson (1989, 2000) and $< z > = 0.182$ as calculated by Radin (2006) from the dream-psi data of Sherwood and Roe (2003), both belonging to the class $p = ½$. Their ratio is 0.28, calling for a correction factor R of about this size to reduce $< z_{n»1} >$ = 0.65 to $< z_{n=1} > = 0.182$. The momentum model with losses compensated by renormalization would almost exactly and that with losses compensated by the null effect more than fulfill this requirement. However, the ratio R differs much less from unity if from the PK data only the quarter of the experiments with the largest values of n is taken into account for which $< z > = 0.41$ (Bösch, Steinkamp, & Boller 2006), and at the same time the result for dream-psi is elevated to the Maimonides value $< z >$ $= 0.26$ for the reasons given in the section **Conjectured Rules**. In order to obtain $< z_{n=1} >$ from $< z_{n»1} >$, the latter is now multiplied by R = 0.75, the correction factor for $\alpha = 1.3$ and $< z > = 0.4066$ applicable in the momentum model with trimmed momenta and in the rectilinear approximation of the field model in the case $p = ½$ (see Table 1). In view of the small difference between $< z_{n=1} > = 0.26$, which is a measured value, and $< z_{n=1} >_{av} = 0.31$, the calculated value for non-existent one-trial PK experiments, one could speak of good agreement between theory and experiment. The agreement is even better for the exact field model with R = 0.66 leading to $< z_{n=1} >_{av}$ $= 0.27$. These considerations suggest that the reduction of the measured $< z >$-score can indeed be explained by the limitations of the hit rate, even without assuming compensated losses. The numbers slightly favor the exact field model over the momentum model with trimmed momenta.

Another comparison concerns the distance between the $< z >$-scores of dream-psi ($p = ½$) and ganzfeld-psi ($p = ¼$). In both cases, the experiments consist of a single trial ($n = 1$). Radin (2006) adopted for dream-psi $(h - p) =$ 0.091 and for ganzfeld-psi $(h-p) = 0.07$, which correspond to $<z> = 0.182$ and 0.162, respectively. This difference could well be explained by the decrease of 15% to be expected if instead of $< z >$ the psi field κ is independent of the chance hit rate p. However, in the preceding paragraph the larger mean-score of dream-psi deduced from the experiments at Maimonides, $< z > =$ 0.26, was preferred. With this value, the effect size of ganzfeld-psi would be 35% less than that of dream-psi. All these z-scores are measured $< z_{n=1} >$ values. For a comparison in terms of many-trial experiments, one has to multiply these values of $< z_{n=1} >$ by the associated inverted correction factors 1/R to obtain $< z_{n»1} >_{re}$, another calculated quantity (the subscript re means reversal). Taking the values of R for $< z > = 0.4066$ from Table 1, one finds that regardless of the model the procedure would add slightly to the size of the discrepancy. I cannot rule out that the large difference is due to an effect-size fluctuation, but consider this to be not very likely. In order to escape a breakdown of Rule 5, let me invoke a possible reason that is the opposite of counting too many trials. The temporal distance of about an hour between ganzfeld experiments may be too small for their reliable separation, and as a consequence two experiments might fuse into a single one whenever the participant does not change. Studies without such a change are mentioned in the literature (see, e.g., Bem & Honorton 1994), but in general there is no attention paid to the problem of separation. Incidentally, Radin's value of the ganzfeld hit rate, $h = 0.32$, is corroborated by other authors. Williams (2011) obtains $h = 0.31$ in his meta-analysis, while Utts, Norris, Suess, and Johnson (2010) and Storm, Tressoldi, and DiRisio (2010) find $h = 0.334$ and 0.32, respectively. The meta-analyses differ by the selection of the data. Radin's value lies in the middle of the others.

The comparisons just made may be regarded as naïve because they neglect the large standard errors and wide confidence intervals linked with the statistical error of the null effect. For instance, with 200 experiments, the 95% confidence interval (two-tailed) of the $< z >$-score remaining after N experiments, $\pm 1.96\sqrt{\frac{p(1-p)}{N}}$, is ±0.06 at $p = ¼$ and ±0.07 at $p = ½$. The scattering of the effect size makes standard error and confidence interval even larger. Without effect-size scattering, the confidence interval of the $< z >$-score of the 450 Maimonides dream-psi experiments, $< z_{n=1} > = 0.26$, would be 0.046.

In the article referred to at the beginning of this section, an attempt was made to explain the extraordinarily high $< z >$-scores found in the ball drawing test of Ertel (2005), a case of ESP with $p = 1/5$, and PK-influenced dice throwing with $p = 1/6$ meta-analyzed by Honorton and Ferrari (1989). They were $< z > = 0.79$ and $< z > = 0.917$, respectively, both measured at large n. For this purpose an ad hoc model was proposed in which $< z >$ diverges as p tends to zero. The rules proposed in the present paper do not permit a divergence of $< z >$, which would be strange anyway. Provided Rule 4 applies, it seems much more likely that the experiments in question were interrupted quite frequently.

Conclusion

In the present paper, six rules presumably or conjecturally holding for statistical psi effects have been formulated. Mainly experiments consisting of hit-or-miss trials, with the chance hit rates $p = ½$ or $p = ¼$, were taken into consideration. The first three rules relate to the absence of dependences on spatial and temporal distances and to the impossibility of markedly increasing the effect size by a collaboration of more than one participant where only one is needed; these three rules appear to be firmly established. The same applies to the fourth rule, but clear specifications of what is needed for the separation of two subsequent experiments with regard to temporal distance and change of participant or participants between experiments are still missing. The fifth rule claims that all statistical psi effects, at least those analyzable in terms of hit rates, are of roughly equal size and that the scattering of the effect size is about as strong as the size itself. The two parts of the rule were combined because the scattering emerges in measurements of the effect size. More data on effect sizes and in particular on effect-size scattering are required for a more precise formulation of this rule and for a final judgement on its validity. There seem to be only three meta-analyses in the literature that present histograms of the z-score distributions (Radin & Nelson 1989, 2000, Schub 2006). The sixth rule is preliminary and speculative, because its claim that these distributions tend to be normal ones is based on the results of essentially the first two of these meta-analyses.

The primary purpose of the calculations was to deal with the possibility of effect sizes per trial being too large to be defined by the mean z-score $< z >$ that like the hit rate h cannot transgress upper and lower limits. Two unlimited effect sizes were derived from different models of the psi effect on condition that they merge with $< z_{\mathrm{n=1}} >$ as $h - p$ approaches zero. In the field model the limited hit rate is a one-to-one function of the unlimited effect size. The deviations of the conjectured true effect size from $< z >$ in this model were estimated from Figure 1 to become substantial in one-trial experiments at hit rate $p = 1/2$ when $< z_{\mathrm{n=1}} >$ rises above 0.5. From Figure 2 it was deduced that they are relatively small in the momentum model up or down to the points where h as a function of the conjectured effect size merges with the lines $h = 1$ or $h = 0$. Any further increase of the effect size beyond these points cannot be detected in the momentum model. A direct check of these predictions to identify the right model is as yet impossible because of the scattering of the effect size and a lack of data suitable for meaningful comparisons. For an indirect check, I included this scattering in the calculations, assuming normal distributions of effect sizes. The first of two comparisons between theory and experiment suggests that the field model is the better choice. The numbers differ too little to allow a final decision. The second comparison deals with a possible dependence of the effect size on p in the field model and seems to lead to a conflict with Rule 5, i.e. the hypothesis of the approximate equality of effect sizes. An experimental reason for this disagreement may be insufficient separation of the ganzfeld-psi experiments resulting in a decrease of their actual number.

The effect sizes of statistical psi are of a magnitude that permits them to noticeably influence the outcome of an experiment, but not large enough to prove the action of psi by a single experiment. A philosophical and at the same time practical reason for this constraint is obvious. Stronger effects could permit us to mentally control (Helfrich 2007) and supervise each other. This would be in conflict with our freedom. We like to believe in the power of good wishes and prayers, but at the same time insist on our autonomy and privacy. In a sense, the uncertainties of psi might reconcile these contradictory demands.

Statistical psi effects are independent of physical laws but do not directly violate them. They utilize the randomness of non-equilibrium states that continually allows choosing among different paths of development. How psi works and the path toward a goal it selects remain mysteries. One may assume that it avoids detours and prefers paths of high probability (which might be another rule). Physics took a long time to recognize the failure of determinism and replace certainties with probabilities. Parapsychology questions this achievement by allowing an influence on randomness. Perhaps parapsychology reveals a bridge between the mind and the material world, in particular the brain. It is astonishing that philosophy and the official churches take hardly any notice of these perspectives. The more rules parapsychology can be shown to obey, the more readily it will be respected by the exact sciences. Therefore, one may hope that it will prove worthwhile to look for rules and with this purpose in mind even to speculate to some extent where the experimental data are still vague or seem to turn out to be incomplete.

Acknowledgments

I would like to express my thanks to Eberhard Bauer of the Institut für Grenzgebiete der Psychologie und Psychohygiene, Freiburg, Germany, for providing me over the years with copies of articles and news in a very patient and supportive way. I am also grateful to Roger Nelson for information and advice, and to a referee for helpful criticism.

Note

1 The scattering of z about $< z > \neq 0$ at the psi-induced hit rate h may be expected to differ from that about $< z > = 0$ at the chance hit rate p, the standard deviation being $\sqrt{h(1-h)n}$ instead of $\sqrt{p(1-p)n}$, an effect that to my knowledge has not yet been observed. The reference state in calculations of deformational energies and chance probabilities of psi-affected z-scores and their distributions is always the ground state with the chance hit rate p.

2 Equation (10) can also be derived by means of a thought experiment with an ideal gas. An ideal gas consisting of n particles (taking the place of trials) in a cylinder of volume V can be divided by a circular impermeable wall into the partial volumes Vh and $V(1 - h)$. Temperature and, in the beginning, pressure are the same in both. Then the wall is shifted

at constant temperature until the partial volumes are Vp and $V(1 - p)$, respectively. The total work of compressing one partial volume (positive) and dilating the other (negative) is

$$E = nk_BT\left[(1-h)\ln\frac{1-h}{1-p} + h\ln\frac{h}{p}\right] \ .$$

This is identical to Equation (10).

References Cited

Baptista, J., Derakshani, M., & Tressoldi, P. E. (2015). Explicit anomalous cognition: A review of the best evidence in ganzfeld, forced-choice, remote viewing, and dream studies. In *Parapsychology: A Handbook for the 21st Century* edited by E. Cardeña, J. Palmer, & D. Marcusson-Claverts, Jefferson, NC: McFarland.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*(3):407–425.

Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, *115*(1):4–18.

Bem, D., Tressoldi, P., Rabeyron, T., & Duggan, M. (2015) Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. *F1000Research* 2015, *4*:1188.

Bösch, H., Steinkamp, F., & Boller, E. (2006a). Examining psychokinesis: The interaction of human intention with random number generators, a meta-analysis. *Psychological Bulletin*, *132*:497–523.

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

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.

Dunne, B. J., & Jahn, R. G. (1992). Experiments in remote human/machine interaction. *Journal of Scientific Exploration*, *6*:311–332.

Dunne, B., & Jahn, R. G. (1995). Consciousness and anomalous physical phenomena. Technical Note 95004. Princeton University, Princeton Engineering Anomalies Research Laboratory. And references cited therein.

Dunne, B. J., & Jahn, R. G. (2003). Information and uncertainty in remote perception research. *Journal of Scientific Exploration*, *17*:207–241.

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

Helfrich, W. (2006). 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. (2011). Are there stable mean values, and relationships between them in statistical parapsychology? *Journal of Scientific Exploration*, *25*:7–28.

Hinterberger, T. (2008). Searching for neuronal markers in psi: Summary of three studies measuring electrophysiology in distant participants. In *Utrecht II: Charting the Future of Parapsychology* edited by C. A. Roe, W. Kramer, and L. Coly, New York: Parapsychology Foundation, pp. 46–62.

Honorton, C., & Ferrari, D. C. (1989). "Future telling": A meta-analysis of forced-choice precognition experiments, 1935–1987. *Journal of Parapsychology*, *53*:281–308.

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

May, E. C. (2015). Experimenter psi: A view of decision augmentation theory. In *Extrasensory Perception, Volume II, Theories of Psi* edited by E. C. May & S. B. Marwaha, Santa Barbara, CA: Praeger, pp. 203–220.

May, E. C., & Marwaha, S. B. (Editors) (2015). *Extrasensory Perception, Volume II, Theories of Psi*. Santa Barbara, CA: Praeger. Including the chapter Entropy and precognition: The physics domain of the multiphasic model of precognition by E. C. May and J. G. Depp, and others by E. C. May et al. on DAT.

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.

May, E. C., Spottiswoode, S. J. P., Utts, J. M., & James, C. L. (1995). Applications of decision augmentation theory. *Journal of Parapsychology*, *59*:221–250.

Mossbridge, J., Tressoldi, P., & Utts, J. (2012). Predictive physiological anticipation preceding seemingly unpredictable stimuli: A meta-analysis. *Frontiers in Psychology*, *3*(October):Article 390.

Nelson, R. (2001). Correlation of global events with REG data: An Internet-based nonlocal anomalies experiment. *The Journal of Parapsychology*, *65*: 247–271.

Nelson, R. D., Radin, D. I., Shoup, R., & Bancel, P. A. (2002). Correlations of continuous random data with major world events. *Foundations of Physics Letters, 15*(6)(December):537–550.

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

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: 1969–2000. Saratoga, CA: Boundary Institute. Online. Print version (2003) Research on mind–matter interaction: Individual intention, in *Healing, Intention and Energy Medicine: Science, Research Methods and Clinical* Implications edited by W. B. Jones & C. C. Crawford, Edinburgh: Churchill Livingston, pp. 39–48.

Radin, D., Nelson, R., Dobyns, Y., & Houtkooper, J. (2006a). Reexamining psychokinesis: Comment on Bösch, Steinkamp and Boller (2006). *Psychological Bulletin, 132*:529–532.

Radin, D., Nelson, R., Dobyns, Y., & Houtkooper, J. (2006b). Assessing the evidence for mind–matter interaction effects. *Journal of Scientific Exploration, 20*:361–374.

Schmidt, S. (2012). Can we help just by good intentions? A meta-analysis of experiments on distant intention effects. *The Journal of Tentative and Complementary Medicine*, *18*:529–539.

Schmidt, S. (2014). *Experimentelle Parapsychologie*. Würzburg: Ergon-Verlag.

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.

Sherwood, S. J., & Roe, C. A. (2003). A review of dream ESP studies conducted since the Maimonides dream ESP studies. In *Psi Wars: Getting to Grips with the Paranormal* edited by J. Alcock, J. Burns, & A. Freeman, Thoverton, UK: Imprint Academic, pp. 85–110.

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

Storm, L., Tressoldi, P. E., & DiRisio, L. (2010) Meta-analysis of free-response studies, 1992–2008: Assessing the noise-reduction model in parapsychology. *Psychological Bulletin*, *136*:471–485.

Tressoldi, P. E., & Utts, J. (2015). Statistical guidelines for empirical studies. In *Parapsychology: A Handbook for the 21st Century* edited by E. Cardeña, J. Palmer, & D. Marcusson-Clavertz, Jefferson, NC: Farland, pp. 83–93.

Utts, J., Norris, M., Suess, E., & Johnson, W. (2010). The strength of evidence versus the power of belief: Are we all Bayesians? Plenary lecture at the Eighth International Conference on Teaching Statistics. In *Data and Context in Statistics Education* edited by C. Reading, Voorburg, The Netherlands: International Statistical Institute.www.stat.auckland.ac.nz/iase

Williams, B. J. (2011). Revisiting the ganzfeld ESP debate: A basic review and assessment. *Journal of Scientific Exploration*, *25*:639–661.