Journal of Scientific Exploration, Vol. 9, No. 4, pp. 453-488, 1995 0892-3310/95

Decision Augmentation Theory: Applications to the Random Number Generator Database

EDWIN C. MAY

Science Applications International Corporation, 330 Cowper St., Suite 200, Palo Alto, CA 94301

JESSICA M. UTTS

University of California, Davis, Division of Statistics, Davis, CA 95616

S. JAMES P. SPOTTISWOODE

Science Applications International Corporation (Consultant), Menlo Park, CA

Abstract — Decision Augmentation Theory (DAT) holds that humans integrate information obtained by anomalous cognition into the usual decision process. The result is that, to a statistical degree, such decisions are biased toward volitional outcomes. We summarize our model and show that the domain over which it is applicable is within a few standard deviations from chance. We contrast the theory's experimental consequences with those of models that treat anomalous effects as due to a force. We derive mathematical expressions for DAT and for force–like models using the normal distribution. The model's predictions for the random number generator database are significantly different for force–like versus informational mechanisms. For large random number generator databases, DAT predicts a zero slope for a least squares fit to a (Z^2, n) scatter diagram, where n is the number of bits resulting from a single run and Z is the resulting Z–score. We find a slope of $(1.73 \pm 3.19) \times 10^{-6}$ ($t = 0.543$, df $= 126$, $p = 0.295$) for the historical binary random number generator database which strongly suggests that some informational mechanism is responsible for the anomaly. In a 2-sequence length analysis of a limited set of data from the Princeton Engineering Anomalies Research laboratory, we find that a force–like explanation misses the observed data by 8.6σ; however, the observed data is within 1.1 o of the DAT prediction. We also apply DAT to one pseudorandom number generator study and find that its predicted slope is not significantly different from the expected value. We provide six circumstantial arguments, which are based upon experimental outcomes against force–like hypotheses. Our anomalous cognition research suggests that the quality of the data is proportional to the total change of Shannon entropy of the target system. We demonstrate that the change of Shannon entropy of a binary sequence from chance is independent of sequence length; thus, we suggest that the change of target entropy may account for successful anomalous cognition and random number generator experiments.

Introduction

We do not have positive definitions of the effects that generally fall under the heading of anomalous mental phenomena.[1] In the crassest of terms, anomalous mental phenomena are what happens when nothing else should, at least as nature is currently understood. In the domain of information acquisition, or anomalous cognition (AC), it is relatively straightforward to design an experimental protocol (Honorton et al., 1990, Hyman and Honorton, 1986) to assure that no known sensory leakage of information can occur. In the domain of macroscopic anomalous perturbation (AP), however, it is often very difficult.

We can divide anomalous perturbation into two categories based on the magnitude of the putative effect. Macro–AP include phenomena that generally do not require sophisticated statistical analysis to tease out weak effects from the data. Examples include inelastic deformations in strain gauge experiments, the obvious bending of metal samples, and a host of possible "field phenomena" such as telekinesis, poltergeist, teleportation, and materialization. Conversely, micro–AP covers experimental data from noisy diodes, radioactive decay and other random sources. These data show small differences from chance expectation and require statistical analysis.

For example, there is now substantial evidence that random number generators, which are designed to produce random binary sequences, deviate from the expected results when a human operator intentionally focuses his or her attention on them. Often the successful experiments are interpreted as a manifestation of some mentally–mediated force. Traditionally this has been called psychokinesis; we call it anomalous perturbation. We are not convinced that a force–like interpretation is correct and propose a different mechanism based upon a mentally–mediated informational process.

One of the consequences of the negative definitions of force–like anomalies is that experimenters must assure that the observables are not due to "known" effects. Traditionally, two techniques have been employed to guard against such interactions:

(1) Complete physical isolation of the target system.
(2) Counterbalanced control and effort periods.

Isolating physical systems from potential "environmental" effects is difficult, even for engineering specialists. It becomes increasingly problematical the more sensitive the macro–AP device. For example Hubbard, Bentley, Pasturel, and Issacs (1987) monitored a large number of sensors of environmental variables that could mimic perturbational effects in an extremely isolated piezoelectric strain gauge. Among these sensors were three–axis accelerometers, calibrated microphones, and electromagnetic and nuclear radiation monitors. In addition, the strain gauges were mounted in a govern-

[1]The Cognitive Sciences Laboratory has adopted the term *anomalous mental phenomena* instead of the more widely known *psi*. Likewise, we use the terms *anomalous cognition* and *anomalous perturbation* for *ESP* and PK, respectively. We have done so because we believe that these terms are more naturally descriptive of the observables and are neutral with regard to mechanisms. These new terms will be used throughout this paper.

ment–approved enclosure to assure no leakage (in or out) of electromagnetic radiation above a given frequency, and the enclosure itself was levitated on an air suspension table. Finally, the entire setup was locked in a controlled access room which was monitored by motion detectors. The system was so sensitive, for example, that it was possible to identify the source of a perturbation of the strain gauge that was due to innocent, gentle knocking on the door of the closed room. The financial and engineering resources to isolate such systems rapidly become prohibitive.

The second method, which is commonly in use, is to isolate the target system within the constraints of the available resources, and then construct protocols that include control and effort periods. Thus, we trade complete isolation for a statistical analysis of the difference between the control and effort periods. The assumption implicit in this approach is that environmental influences of the target device will be random and uniformly distributed in both the control and effort conditions, while anomalous effects will tend to occur in the effort periods. Our arguments in favor of an anomaly, then, are based on statistical inference and we must consider, in detail, the consequences of such analyses.

Background

As the evidence for anomalous mental phenomena becomes more widely accepted (Bem and Honorton, 1994; Utts, 1991; Radin and Nelson, 1989) it is imperative to determine their underlying mechanisms. Clearly, we are not the first to begin thinking of potential models. In the process of amassing incontrovertible evidence of an anomaly, many theoretical approaches have been examined; in this section we outline a few of them. It is beyond the scope of this paper, however, to provide an exhaustive review of the theoretical models; a good reference to an up–to–date and detailed presentation is Stokes (1987).

Brief Review of Models

Two fundamentally different types of models of anomalous mental phenomena have been developed: those that attempt to order and structure the raw observations in experiments (i.e., phenomenological models), and those that attempt to explain these phenomena in terms of modifications to existing physical theories (i.e., fundamental models). In the history of the physical sciences, phenomenological models, such as the Snell's law of refraction or Ampere's law for the magnetic field due to a current, have nearly always preceded fundamental models, such as quantum electrodynamics and Maxwell's theory. In producing useful models of anomalies it may well be advantageous to start with phenomenological models, of which DAT is an example.

Psychologists have contributed interesting phenomenological approaches. Stanford (1974a and 1974b) proposed PSI–Mediated Instrumental Response (PMIR). PMIR states that an organism uses anomalous mental phenomena to optimize its environment. For example, in one of Stanford's classic experiments (Stanford, Zenhausern, Taylor, and Dwyer, 1975) subjects were offered a covert opportunity to stop a boring task prematurely if they exhibited unconscious anomalous perturbation by perturbing a hidden random number generator. Overall, the experiment was significant in the unconscious tasks; it was as if the participants were unconsciously scanning the extended environment for any way to provide a more optimal situation than participating in a boring psychological task!

As an example of a fundamental model, Walker (1984) proposed a literal interpretation of quantum mechanics and posited that since superposition of eigenstates holds, even for macrosystems, anomalous mental phenomena might be due to macroscopic examples of quantum effects. These ideas spawned a class of theories, the so–called observation theories, that were either based upon quantum formalism conceptually or directly (Stokes, 1987). Jahn and Dunne (1986) have offered a "quantum metaphor" which illustrates many parallels between these anomalies and known quantum effects. Unfortunately, these models either have free parameters with unknown values, or are merely metaphors. Some of these models propose questionable extensions to existing theories. For example, even though Walker's interpretation of quantum mechanical formalism might suggest wave–like properties of macrosystems, the physics data to date not only show no indication of such phenomena at room temperature but provide considerable evidence to suggest that macrosystems lose their quantum coherence above 0.5 Kelvins (Washburn and Webb, 1986) and no longer exhibit quantum wave–like behavior.

This is not to say that a comprehensive model of anomalous mental phenomena may not eventually require quantum mechanics as part of its explanation, but it is currently premature to consider such models as more than interesting speculation. The burden of proof is on the theorist to show why systems, which are normally considered classical (e.g., a human brain), are, indeed, quantum mechanical. That is, what are the experimental consequences of a quantum mechanical system over a classical one?

Our Decision Augmentation Theory is phenomenological and is a logical and formal extension of Stanford's elegant PMIR model. In the same manner as early models of the behavior of gases, acoustics, or optics, DAT tries to subsume a large range of experimental measurements into a coherent lawful scheme. Hopefully this process will lead the way to the uncovering of deeper mechanisms. In fact DAT leads to the idea that there may be only one underlying mechanism of all anomalous mental phenomena, namely a transfer of information between events separated by negative time intervals.

Historical Evolution of Decision Augmentation

May, Humphrey, and Hubbard (1980) conducted a careful random number generator (RNG) experiment which was distinguished by the extreme engineering and methodological care that was taken to isolate any potentially known physical interactions with the source of randomness (D. Druckman and J. A. Swets, page 189, 1988). It is beyond the scope of this paper to describe this experiment completely; however, those specific details which led to the idea of Decision Augmentation are important for the sake of historical completeness. The authors were satisfied that they had observed a genuine statistical anomaly and additionally, because they had developed an accurate mathematical model of the random device, they were assured that the deviations were not due to any known physical interactions. They concluded, in their report, that some form of anomalous data selection had occurred and named it ***Psychoenergetic Data Selection.***

Following a suggestion by Dr. David R. Saunders of MARS Measurement and Associates, we noticed in 1986 that the effect size in binary RNG studies varied on the average as one over the square root of the number of bits in the sequence. This observation led to the development of the ***Intuitive Data Sorting*** model that appeared to describe the RNG data to that date (May, Radin, Hubbard, Humphrey, and Utts, 1985). The remainder of this paper describes the next step in the evolution of the theory which is now named ***Decision Augmentation Theory.***

Decision Augmentation Theory — A General Description

Since the case for AC–mediated information transfer is now well established (Bem and Honorton, 1994) it would be exceptional if we did ***not*** integrate this form of information gathering into the decision process. For example, we routinely use real–time data gathering and historical information to assist in the decision process. Why, then, should we not include AC in the decision process? DAT holds that AC information is included along with the usual inputs that result in a final human decision that favors a "desired" outcome. In statistical parlance, DAT says that a slight, systematic bias is introduced into the decision process by AC.

This philosophical concept has the advantage of being quite general. To illustrate the point, we describe how the "cosmos" determines the outcome of a well–designed, hypothetical experiment. To determine the sequencing of conditions in an RNG experiment, suppose that the entry point into a table of random numbers will be chosen by the square root of the barometric pressure as stated in the weather report that will be published seven days hence in the ***New York Times.*** Since humans are notoriously bad at predicting or controlling the weather, this entry point might seem independent of a human decision; but why did we "choose" seven days in advance? Why not six or eight? Why the ***New York Times*** and not the ***London Times?*** DAT would suggest that the selection of seven days, the ***New York Times,*** the barometric pressure, and square root function were better choices, either individually or collectively, and that other decisions would not have led to as significant an outcome. Other non–technical decisions may also be biased by AC in accordance with DAT. When should we schedule a Ganzfeld session; who should be the experimenter in a series; how should we determine a specific order in a tri–polar protocol? DAT explains anomalous mental phenomena as a process of judicious Sampling from a world of events that are unperturbed. In contrast, force–like models hold that some kind of mentally–mediated force perturbs the world. As we will show below, these two types of models lead to quite different predictions.

It is important to understand the domain in which a model is applicable. For example, Newton's laws are sufficient to describe the dynamics of mechanical objects in the domain where the velocities are very much smaller than the speed of light, and where the quantum wavelength of the object is very small compared to the physical extent of the object. If these conditions are violated, then different models must be invoked (e.g., relativity and quantum mechanics, respectively). The domain in which DAT is applicable is when experimental outcomes are in a statistical regime (i.e., a few standard deviations from chance). In other words, could the measured effect occur under the null hypothesis? This is not a sharp–edged requirement but DAT becomes less apropos the more a single measurement deviates from mean–chance–expectation (MCE). We would not invoke DAT, for example, as an explanation of levitation if one found the authors hovering near the ceiling! The source of the statistical variation is unrestricted and may be of classical or quantum origin, because a potential underlying mechanism for DAT is precognition. By this means, experiment participants become statistical opportunists.

Development of a Formal Model

While DAT may have implications for anomalous mental phenomena in general, we develop the model in the framework of understanding experimental results. In particular, we consider anomalous perturbation versus anomalous cognition in the form of decision augmentation in those experiments whose outcomes are in the few–sigma, statistical regime.

We define four possible mechanisms for the results in such experiments:

1) ***Mean Chance Expectation.*** The results are at chance. That is, the deviation of the dependent variable meets accepted criteria for MCE. In statistical terms, we have measurements from an ***unperturbed*** parent distribution with ***unbiased*** sampling.

2) ***Anomalous Perturbation.*** Nature is modified by some anomalous interaction. That is, we expect an interaction of a "force" type. In statistical parlance, we have measurements from a ***perturbed*** parent distribution with ***unbiased*** sampling.

3) ***Decision Augmentation.*** Nature is unchanged but the measurements are biased. That is, AC information has "distorted" the sampling. In statistical terms, we have measurements from an ***unperturbed*** parent distribution with ***biased*** sampling.

4) ***Combination.*** Nature is modified and the measurements are biased. That is, both anomalous effects are present. In statistical parlance, we have conducted ***biased*** sampling from ***aperturbed*** parent distribution.

There may be other explanations of the deviations such as unintentional or intentional selection of data or results. A comprehensive critique of many such interpretations can be found in Radin and Nelson (1989), and since they provide convincing arguments against these interpretations, we only consider the first three here.

General Considerations and Definitions

Since the formal discussion of DAT is statistical, we will describe the overall context for the development of the model from that perspective. Consider a random variable, X, that can take on continuous values (e.g., the normal distribution) or discrete values (e.g., the binomial distribution). Examples of X might be the hit rate in an RNG experiment, the swimming velocity of single cells, or the mutation rate of bacteria. Let Y be the average of X computed over ***n*** values, where ***n*** is the number of items that are collected as the result of a single decision — one trial. Often this may be equivalent to a single effort period, but it also may include repeated efforts. The key point is that, regardless of the effort style, the average value of the dependent variable is computed over the ***n*** values resulting from one decision point. In the examples above, ***n*** is the sequence length of a single run in an RNG experiment, the number of swimming cells measured during the trial, or the number of bacteria–containing test tubes present during the trial. As we will show below, force–like effects require that the Z–score, which is computed from the Ys, increase as the square root of ***n.*** In contrast, informational effects will be shown to be independent of ***n.***

Assumptions for DAT

We assume that the parent distribution of a physical system remains ***unperturbed;*** however, the measurements of the physical system are systematically biased by some AC–mediated informational process.

Since the deviations seen in experiments in the statistical regime tend to be small in magnitude, it is safe to assume that the measurement biases will also be small; therefore, we assume small shifts of the mean and variance of the sampling distribution. Figure 1 shows the distributions for biased and unbiased measurements.

The biased sampling distribution shown in Figure 1 is assumed to be normally distributed as:

$$Z \sim N(\mu_z, \sigma_z^2)$$

where μ_z and σ_z are the mean and standard deviation of the sampling distribution.

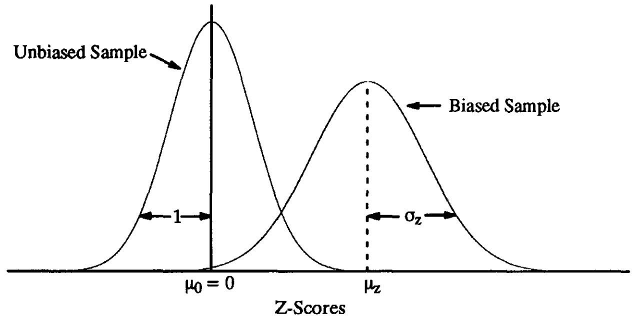

Fig. 1. Sampling distribution under DAT.

Assumptions for an Anomalous Perturbation Model

DAT can be contrasted to force–like effects. With a few exceptions reported in the literature of "field" phenomena, anomalous perturbation appears to be relatively "small." Thus, we begin with the assumption that a putative anomalous force would give rise to a perturbational interaction, by which we mean that, given an ensemble of entities (e.g., binary bits, cells), an anomalous force would act equally on each member of the ensemble, on the average. We call this type of interaction micro–AP.

Figure 2 shows a schematic representation of probability density functions for a parent distribution under the micro–AP assumption and an unperturbed parent distribution. In the simplest micro–AP model, the perturbation induces a change in the mean of the parent distribution but does not effects its variance. We parameterize the mean shift in terms of a multiplier of the initial standard deviation. Thus, we define an AP–effect size as:

$$\varepsilon_{AP} = \frac{(\mu_1 - \mu_0)}{\sigma_0},$$

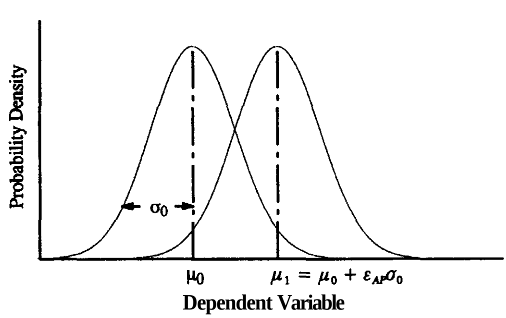

Fig. 2. Parent distribution for micro-AP.

where μ_1 and are μ_0 the means of the perturbed and unperturbed distributions, respectively, and where σ_0 is the standard deviation of the unperturbed distribution.

For the moment, we consider ε_{AP} as a parameter which, in principle, could be a function of a variety of variables (e.g., psychological, physical, environmental, methodological). As we develop DAT for specific distributions and experiments, we will discuss this functionality of ε_{AP}.

Calculation of $E(Z^2)$

We compute the expected value and variance of Z^2 for mean chance expectation and under the force–like and information assumptions. We do this for the normal and binomial distributions. The details of the calculations can be found in the Appendix; however, we summarize the results in this section. Table 1 shows the results assuming that the parent distribution is normal.[2]

We wish to emphasize at this point that in the development of the mathematical model, the parameter ε_{AP} for micro-AP, and the parameters μ_z and σ_z in DAT may all possibly depend upon n; however, for the moment, we *assume* that they are all n–independent. We shall discuss the consequences of this assumption below.

Figure **3** displays these theoretical calculations for the three mechanisms graphically.

This formulation predicts grossly different outcomes for these models and, therefore, is ultimately capable of separating them, even for very small effects. The important differences are in the slope and intercept values. MCE gives a slope of zero and an intercept of one. DAT predicts a slope of zero, but an intercept greater than one, and Micro–AP predicts an intercept of one, but a slope greater than zero.

TABLE 1
Normal Parent Distribution

	Mechanism		
Quantity	MCE	Micro-AP	DAT
$E(Z^2)$	1	$1+\varepsilon_{AP}^2 n$	$\mu_z^2+\sigma_z^2$
$Var(Z^2)$	2	$2(1+2\varepsilon_{AP}^2 n)$	$2(\sigma_z^4+2\mu_z^2\sigma_z^2)$

Monte Carlo Verification

The expressions shown in Table 1 are representations which arise from simple algebraic manipulations of the basic mathematical assumptions of the models. To verify that these expressions give the expected results, we used a

[2]For completeness, the appendix also shows the same calculations for the binomial distribution. Since the normal approximation to the binomial distribution is valid for the RNG database, we only discuss the normal formalism in the body of this paper.

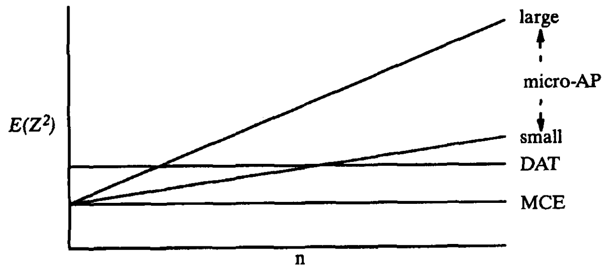

Fig. 3. Predictions of MCE, micro-AP, and DAT.

published pseudo random number generator (Lewis, 1975) with well–understood properties to produce data that mimicked the results under three models (i.e., MCE, micro–AP and DAT). Our standard implementation of the pseudo–RNG allows the integers in the range $(0, 2^{15}-1)$ as potential seeds. For the sequence lengths 100,500,1000, and 5000, we computed Z–scores for all possible seeds with an effect size of 0.0 to simulate MCE and an effect size of 0.03 to simulate micro–AP. To simulate DAT, we used the fact that in the special case where the effect size varies as $1/n^{1/2}$, micro–AP and DAT are equivalent. For this case we used effect sizes of 0.030, 0.0134, 0.0095, and 0.0042 for the above sequence lengths, respectively. Figures 4a–c show the results of 100 trials, which were chosen randomly from the appropriate Z–score data sets, at each of the sequence lengths for each of the models. In each Figure, MCE is indicated by a horizontal solid line at $Z^2 = 1$

The slope of a least squares fit computed under the MCE simulation was $(-2.81 \pm 2.49) \times 10^{-6}$, which corresponded to a p–value of 0.812 when tested against zero, and the intercept was 1.007 ± 0.005, which corresponds to a p–value of 0.131 when tested against one. Under the micro–AP model, an estimate of the effect size using the expression in Table 1 was $\varepsilon_{AP} = 0.0288 \pm 0.002$, which is in good agreement with 0.03, the value that was used to create the data. Similarly, under DAT the slope was $(-2.44 \pm 57.10) \times 10^{-6}$, which corresponded to a p–value of 0.515 when tested against zero, and the intercept was 1.050 ± 0.001, which corresponds to a p–value of 2.4×10^{-4} when tested against one.

Thus, we are able to say that the Monte Carlo simulations confirm the simple formulation shown in Table 1.

Retrospective Tests

It is possible to apply DAT retrospectively to any body of data that meet certain constraints. It is critical to keep in mind the meaning of n — the number of measures of the dependent variable over which to compute an average dur-

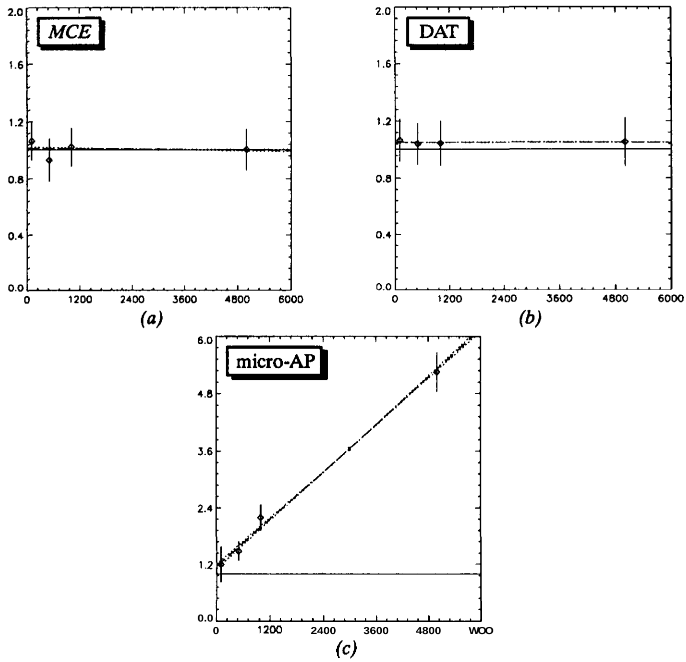

Figure 4. Z^2 vs n for Monte Carlo Simulations of *MCE*, micro-AP, and DAT.

ing a single trial following a single decision. In terms of their predictions for experimental results, the crucial distinction between DAT and the micro–AP model is the dependence of the results upon *n*; therefore, experiments which are used to test these theories ideally should be those in which experiment participants are blind to *n*, and where the distribution of *n* does not contain extreme outliers.

Aside from these considerations, the application of DAT is straight forward. Having identified the unit of analysis and *n*, simply create a scatter diagram of points (Z^2, n) and compute a weighted least square fit to a straight line. Table 1 shows that for the micro–AP model, the slope of the resulting fit is the square of the AP–effect size. A Student's t–test may be used to test the hypothesis that the AP–effect size is zero, and thus test for the validity of the micro–AP model. If the slope is zero, these same tables show that the intercept may be interpreted as a strength parameter for DAT. In other words, an intercept larger than one would support the DAT model, while a slope greater than zero would support the micro–AP model.

Historical Binary RNG Database

Radin and Nelson (1989) analyzed the complete literature (i.e., over 800 individual studies) of consciousness–related anomalies in random physical systems. They demonstrated that a robust statistical anomaly exists in that database. Although they analyzed this data from a number of perspectives, they report an average $Z/n^{1/2}$ effect size of approximately 3×10^{-4}, regardless of the analysis type. Radin and Nelson did not report p–values, but they quote a mean Z of 0.645 and a standard deviation of 1.601 for 597 studies. We compute a single–mean t–score of 9.844, df = 596 @ = 3.7×10^{-23}).

We returned to the original publications of all the binary RNG studies from those listed by Radin and Nelson and identified 128 studies in which we could compute, or were given, the average Z–score, the number of runs, N, and the sequence length, n, which ranged from 16 to 10,000. For each of these studies we computed:

$$\overline{Z^2} = \mu_z^2 + \left(\frac{N-1}{N}\right)s_z^2 .$$

Since we were unable to determine the standard deviations of the Z–scores from the literature, we assumed that $s_z = 1.0$ for each study. We see from Table 1 that under mean chance expectation the expected variance of each Z^2 is 2.0 so that the estimated standard deviation for the Z^2 for a given study is $(2.0/N)^{1/2}$.

Figure 5 shows a portion of the 128 data points (Z^2, n). MCE is shown as a solid line (i.e., $Z^2 = 1$), and the expected best–fit lines for two assumed AP effect sizes of 0.01 and 0.003, respectively, are shown as short dashed lines. We calculated a weighted (i.e., using $N/2.0$ as the weights) least squares fit to an a + bn straight line for the 128 data points and display it as a long–dashed line. For clarity, we have offset and limited the Z^2 axis and have not shown the error bars for the individual points, but the weights and all the data were used in the least squares fit. We found an intercept of a = 1.036±0.004. The 1-σ standard error for the intercept is small and is shown in Figure 5 in the center of the sequence range. The t–score for the intercept being different from 1.0 (i.e., t = 9.1, df = 126, p = 4.8×10^{-20}) is in good agreement with that derived from Radin and Nelson's analysis. Since we set standard deviations for all the Z's equal to one; and since Radin and Nelson showed that the overall standard deviation was 1.6, we would expect that our analysis would be more conservative than theirs because a larger standard deviation would increase our computed value for the intercept.

The important result, however, was that the slope of the best–fit line was b = $(1.73 \pm 3.19) \times 10^{-6}$ (t = 0.543, df = 126, p = 0.295), which is not significantly different from zero. Adding and subtracting one standard error to the slope estimate produces an interval that encompasses zero. Even though a very small AP effect size might fit the data at large sequence lengths, it is clear in Figure 5 what happens at small sequence lengths; an $\varepsilon_{AP} = 0.003$, suggests a linear fit that is significantly below the actual fit.

The sequence lengths from this database are not symmetric nor are they uniformly distributed; they contain outliers (i.e., median = 64, average = 566).

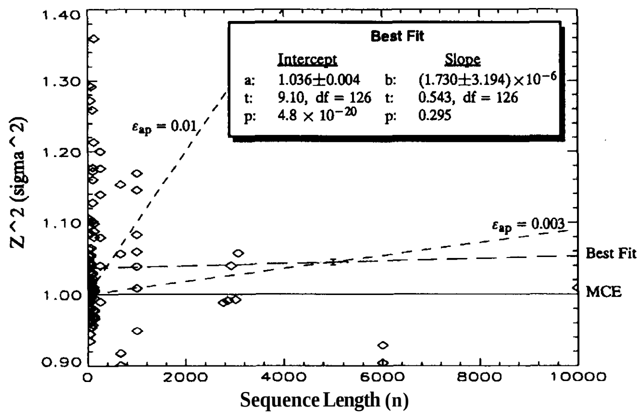

Figure 5. Binary RNG Database: Slope and Intercept for Best Fit Line.

Figure 6 shows that the lower half of the data, however, is symmetric and nearly uniformly distributed (i.e., median = 35, average = 34). Since the criterion for a valid retrospective test is that n should be uniform, or at least not contain outliers, we analyzed the two median halves independently. The intercept for the weighted best–fit line for the uniform lower half is a = 1.022±0.006 (t = 3.63, df = 62, p = 2.9 x 10^{-4}), and the slope is b = (−0.03423.70) x 10^{-4} (t = −0.010, df = 62, p = 0.504). The fits for the upper half yield a = 1.064±0.005 (t = 13.47, df = 62, p = 1.2 × 10^{-41}) and b = (−4.52±2.38) x 10^{-6} (t = −1.903, df = 62, p = 0.969), for the intercept and slope, respectively.

Since the best retrospective test for DAT is one in which the distribution of n contains no outliers, the statistically zero slope for the fit to the lower half of the data is inconsistent with a simple AP model. Although the same conclusion could be reached from the fits to the database in its entirety (i.e., Figure 5), we suggest caution in that this fit could possibly be distorted by the distribution of the sequence lengths. That is, a few points at large sequence lengths can easily influence the slope. Since the slope for the upper half of the data is statistically slightly negative, it is problematical to assign an imaginary AP effect size to these data. More likely, the results are distorted by a few outliers in the upper half of the data.

From these analyses, it appears that Z^2 does not linearly depend upon the sequence length; however, since the scatter is so large, even a linear model is not a good fit (i.e., $\chi^2 = 171.2$, $df = 125$, $p = 0.0038$), where χ^2 is a goodness–of–fit measure in general given by:

$$\chi^2 = \sum_{j=1}^{\nu} \frac{1}{\sigma_j^2}(y_j - f_j)^2$$

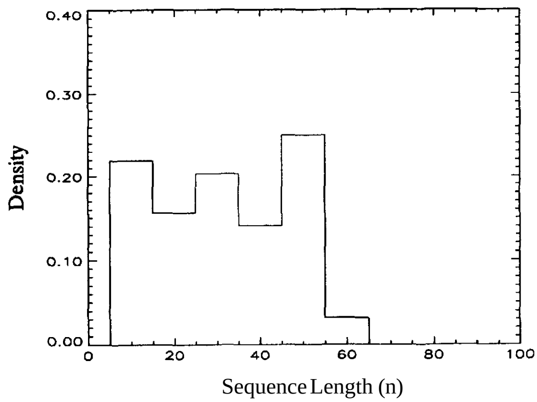

Fig. 6. Historical Database: Distribution of Sequence Lengths < 64.

where the σ_j are the errors associated with data point y_j, f_j is the value of the fitted function at point j, and v is the number of data points.

A "good" fit to a set of data should lead to a non–significant χ^2. The fit is not improved by using higher order polynomials (i.e., $\chi^2 = 170.8$, df = 124; $\chi^2 = 174.1$, df = 123; for quadratics and cubics, respectively). If, however, the AP effect size was any monotonic function of n other than the degenerate case where the AP effect size is proportional to $1/n^{1/2}$, it would manifest as a non–zero slope in the regression analysis.

Within the limits of this retrospective analysis, we conclude for RNG experiments that we must reject all influence models which propose a shift of the mean of the parent distribution.

Princeton Engineering Anomalies Research Laboratory RNG Data

The historical database that we analyzed does not include the extensive RNG data from the Princeton Engineering Anomalies Research (PEAR) laboratory since the total number of bits in their experiments exceeds the total amount in the entire historical database. For example, in a recent report Nelson, Dobyns, Dunne, and Jahn (1991) analyze 5.6 x 10^6 trials all at $n = 200$ bits. In this section, we apply DAT retrospectively to their published work where they have examined other sequence lengths; however, even in these cases, they report over five times as much data as in the historical database.

Jahn (1982) reported an initial RNG data set with a single operator at n = 200 and 2,000. Data were collected both in the automatic mode (i.e., a single button press produced 50 trials at *n)* and the manual mode (i.e., a single button press produced one trial at n). From a DAT perspective, data were actually collected at four values of n (i.e., 200, 2000, 200 x 50 = 10,000, and 2000 × 50 = 100,000). Unfortunately data from these two modes were grouped together and reported only at 200 and 2,000 bit/trial. It would seem, therefore, we would be unable to apply DAT to these data. Jahn, however, reports that the different modes "...give little indication of importance of such factors in the overall performance." This qualitative statement suggests that the micro–AP model is indeed not a good description for these data, because, under micro–AP, we would expect stronger effects (i.e., higher Z-scores) at the longer sequence lengths.

Nelson, Jahn, and Dunne (1986) describe an extensive RNG and pseudo–RNG database in the manual mode only (i.e., over 7×10^6 trials); however, whereas Jahn provides the mean and standard deviations for the hits, Nelson et *al.* report only the means. We are unable to apply DAT to these data, because any assumption about the standard deviations would be highly amplified by the massive data set.

As part of a cooperative agreement in 1987 between PEAR and the Cognitive Sciences Program at SRI International, we analyzed a set of RNG data from a single operator. Since they supplied the raw data for each button press, we were able to analyze this data at two extreme values of n. We combined the individual trial Z–scores for the high and low aims, because our analysis is two–tailed, in that we examine Z^2.

Given that the data sets at $n = 200$ and 100,000 were independently significant (Stouffer's Z of 3.37 and 2.45, respectively), and given the wide separation between the sequence lengths, we used DAT as a retrospective test on these two data points.

Because we are examining only two values of n, we do not compute a best–fit slope. Instead, as outlined in May, Utts, and Spottiswoode (1995), we compare the micro–AP prediction to the actual data at a single value of n.

At n = 200, 5918 trials yielded $Z = 0.044 \pm 1.030$ and $Z^2 = 1.063 \pm 0.019$. We compute a proposed AP effect size $Z/200^{1/2} = 3.10 \times 10^{-3}$. With this effect size, we computed what would be expected under the micro–AP model at n = 100,000. Using the theoretical expressions in Table 1, we computed $Z^2 = 1.961 \pm 0.099$. The 1 σ error is derived from the theoretical variance divided by the actual number of trials (597) at n = 100,000. The observed values were $Z = 0.100 \pm 0.997$ and $Z^2 = 1.002 \pm 0.050$. A t–test between the observed and expected values of Z^2 gives $t = 8.643$, df = 1192. Considering this t as equivalent to a Z, the data at $n = 100{,}000$ fails to meet what would be expected under the influence model by 8.6σ. Suppose, however, that the effect size observed at n = 100,000 (3.18×10^{-4}) better represents the AP effect size. We computed the predicted value of $Z^2 = 1.00002 \pm 0.018$ for n = 200. Using a t–test for the difference between the observed value and this predicted one gives t = 2.398, df= 11,834. The micro–AP model fails in this direction by more than 2.3σ. DAT predicts that Z^2 would be statistically equivalent at the two sequence lengths, and we find that to be the case (t = 1.14, df = 6513, p = 0.127).

Jahn (1982) indicates in their RNG data that "Traced back to the elemental binary samples, these values imply directed inversion from chance behavior of about one or one and a half bits in every one thousand... ." Assuming 1.5 excess bits/1000, we calculate an AP effect size of 0.003, which is consistent with the observed value in their n = 200 data set. Since this was the value we used in our DAT analysis, we are forced to conclude that this data set from PEAR is inconsistent with the simple micro–AP model, and that Jahn's statement is not a good description of the anomaly.

We urge caution in interpreting these calculations. As is often the case in a retrospective analysis, some of the required criteria for a meaningful test are violated. These data were not collected when the operators were blind to the sequence length. Secondly, these data represent only a fraction of PEAR's database.

A Prospective Test of DAT

In developing a methodology for future tests, Radin and May (1986) worked with two operators who had previously demonstrated strong ability in RNG studies. They used a pseudo–RNG, which was based on a shift–register algorithm by Kendell and has been shown to meet the general criteria for "randomness" (Lewis, 1975), to create the binary sequences.

The operators were blind to which of nine different sequences (i.e., n = 101, 201, 401, 701, 1001, 2001, 4001, 7001, 10001 bits)[3] were used in any given trial, and the program was such that the trials lasted for a fixed time period and feedback was presented only after the trial was complete. Thus, the criteria for a valid test of DAT had been met, except that the source of the binary bits was a pseudo–RNG.

We re–analyzed the combined data from this experiment with the current *Z*–score formalism of DAT. For the 200 individual runs (i.e. 10 at each of the sequence lengths for each of the two participants) we found the best fit line to yield a slope = $4.3 \times 10^{-8} \pm 1.6 \times 10^{-6}$ ($t = 0.028$, $df = 8$, $p = 0.489$) and an intercept = $1.16 + 0.06$ ($t = 2.89$, $df = 8$, $p = 0.01$). The slope interval easily encompasses zero and is not significantly different from zero, the intercept significance level ($p = 0.01$) is consistent with what Radin and May reported earlier.

Since the pseudo–RNG seeds and bit streams were saved for each trial, it was possible to determine if the experiment sequences exactly matched the ones produced by the shift register algorithm; they did. Since their UNIX–based Sun Microsystems workstations were synchronized to the system clock, any momentary interruption of the clock would "crash" the machine, but no such crashes occurred. Therefore, we believe no force–like interaction occurred.

To explore the timing aspects of the experiment Radin and May reran each run with pseudo–RNG seeds ranging from –5 to +5 clock ticks (i.e., 20 ms/tick) from the actual seed used in the run. We plot the resulting run effect

[3]The original IDS analysis required the sequence lengths to be odd because of the logarithmic formalism.

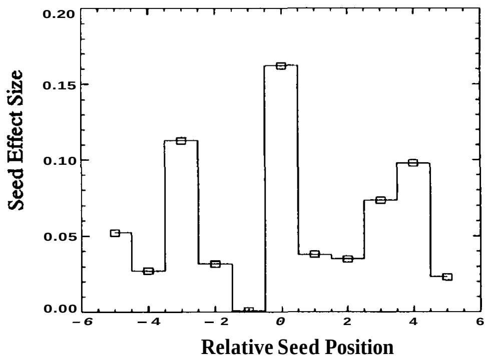

Fig. 7 Seed Timing for Operator 531 — 298 Runs.

sizes, which we computed from the experimental F–ratios (Rosenthal, 1991), for operator 531 in Figure 7. The estimated standard errors are the same for each seed shift and equal 0.057.

Radin and May erroneously concluded that the significant differences between zero and adjacent seed positions was meaningful, and that the DAT ability was effective within 20 milliseconds. In fact, the situation shown in Figure 7 is ***expected.*** Differing from true random number generators in which slight changes in entry points produce essentially the same sequence, pseudo–RNGs produced totally different sequences as a function of single digit seed changes. Thus, it would be surprising if the seed–shift display produced anything but a spike at seed shift zero.

Walker (1987) proposed that individuals would have to exhibit a physiologically impossible control over timing (e.g., when to press a button). As evidence apparently in favor of such an exquisite timing ability, he referred to the data presented by Radin and May (1986) that we have discussed above. Walker suggested that Radin and May's result, therefore, supported his quantum mechanical observer theory. It is beyond the scope of this paper to critique Walker's quantum mechanical models, but we would hope they do *not* depend upon his analysis of Radin and May's results since the result we show in Figure 7 is expected and does not represent the precision of the operator's reaction time.

We must consider how it is possible with normal human reactions to obtain significant scores, which can only happen in 20 ms windows. In typical visual reaction time measurements, Woodworth and Schlosberg (1960) found a standard deviation of 30 ms. If we assume these human reactions are typical of those for AC performance and are normally distributed, we compute the ***maximum*** probability of being within a 20 ms window, which is centered about the mean, of 23.5%. For the worst case, the operators must "hit" significant seeds less often than 23.5% of the time. Radin and May do not report the number of significant runs, so we provide a worst–case estimate. Given that they quote a p–value of 0.005 for 500 trials, we find that 39 trials must be independently significant. That is, the accumulated binomial probability is 0.005 for 39 hits in 500 trials with an event probability of 0.05. This corresponds to a hitting rate (i.e., 391500) of only 7.8%, a value well within the capability of human reaction times. We recognize that it is not a requirement to hit only on significant seeds; however, all other seeds leading to positive Z–scores are less restrictive than the case we have presented.

The zero–center "spike" in Figure 7 misled Walker and others into thinking that exceptional timing was required to produce the observed deviations. As we have shown this is not the case, and, therefore, Walker's criticism of the theory is not valid.

From this prospective test of DAT, we conclude that for pseudo–RNGs it is possible to select a proper entry point into a bit stream, even when the seed intervals are as small as 20 ms, to produce significant deviations from mean chance expectation. These deviations are independent of sequence length.

Critical Review of Dobyns — 1993

Dobyns (1993) presents a method for comparing what he calls the "influence" and "selection" models, corresponding to what we have been calling DAT and micro–AP. He uses data from 490 "tripolar sets" of experimental runs at PEAR. For each set, there was a high aim, a baseline and a low aim condition.

The three values produced were then sorted into which one was actually highest, in the middle, and lowest for each set. The data were then summarized into a 3×3 matrix, where the rows represented the three intentions, and the columns represented the actual ordering. If every attempt had been successful, the diagonal of the matrix would consist of the number of tripolar sets, namely

TABLE 2
Scoring data from (Dobyns, 1993).

	Intention			
Actual	**High**	**Middle**	**Low**	**Total**
High	180	167	143	490
Baseline	159	156	175	490
Low	151	167	172	490
Total	490	490	490	

490. We present the data portion of Dobyns' Table from page 264 of the reference as our Table 2:

Dobyns computes an aggregate likelihood ratio of his predictions for the DAT and micro–AP models and concludes in favor the the influence model with a ratio of 28.9 to one.

However, there are serious problems with the methods used in Dobyns' paper. In this paper we outline only two of the difficulties. To fully explain them would require a level of technical discussion not suitable for a short summary such as this.

One problem is in the calculation of the likelihood ratio function using his Equation 6, which we reproduce from page 265 of the reference:

$$B(p/q) = \frac{p_1^{n_1} p_2^{n_2} p_3^{n_3}}{q_1^{n_1} q_2^{n_2} q_3^{n_3}} = \left[\frac{p_1}{q_1}\right]^{n_1} \left[\frac{p_2}{q_2}\right]^{n_2} \left[\frac{p_3}{q_3}\right]^{n_3},$$

where p and q are the predicted rank frequencies for each aim under the influence and selection models, respectively, and the n are the observed frequencies for each aim. We agree that this relationship correctly gives the likelihood ratio for comparing the two models for one row of Table 2. However, immediately following that equation, Dobyns writes, "The aggregate likelihood of the hypothesis over all three intentions may be calculated by repeating the individual likelihood calculation for each intention, and the total likelihood will simply be the product of factors such as (6) above for each of the three intentions."

That statement is incorrect. A combined likelihood is found by multiplying the individual likelihoods only if the random variables are independent of each other (DeGroot, 1986, p. 145). Clearly, the rows of the table are not independent. In fact, if you know any two of the rows, the third is determined exactly. The correct likelihood ratio needs to build that dependence into the formula.[4]

A second technical problem with the conclusion that the data support the influence model is that the method itself strongly supports the influence model. As noted by Dobyns, "In fact, applying the test to data sets that, by construction, contain no effect, yields strong odds (ranging, in a modest Monte Carlo database, from 8.5 to over 100) in favor of the influence model (page 268)." The actual data in his paper yielded odds of 28.9 to one in favor of the influence model; however, this value is well within the reported limits from his "influence–less" Monte Carlo data.

Under DAT, it is possible that AC–mediated selection might occur at the protocol level, but the primary way is through timing — initiating a run to capitalize upon a locally deviant subsequence. How this might work in dynamic RNG devices is clear; wait until such a deviant sequence is in your immediate future and initiate the run in time to capture it. With "static" devices, such as PEAR's random mechanical cascade device, how timing enters in is less obvious. Under closer inspection, however, even with this device there is a statistical variation among unattended control runs. That is, there is never a series of control runs that give exactly the same mean. Physical effects, such as Brownian motion, temperature gradients, etc., can account for the observed variance in the absence of human operators. Thus, ***when*** a run is initiated to capture fa-

[4]Dobyns agrees on this point — private communication.

vorable local "environmental" factors, even for "static" devices, remains the operative issue with regard to DAT. Dobyns does not consider this case at all in his analysis. If DAT enters in at the protocol selection, as it probably does, it is likely to be a second–order contribution because of the limited possibilities from which to select (i.e., six in the tripolar case).

Finally, a major problem with Dobyns' conclusion, which was pointed out when he first presented this paper at a conference (May, 1990), is that the likelihood ratio supports the influence model even for their pseudo–RNG data. Dobyns dismisses this finding (page 268) all too easily given the preponderance of evidence that suggest that no influence occurs during pseudo–RNG studies (Radin and May, 1986).

Aside from the technical flaws in Dobyns' likelihood ratio arguments, and even ignoring the problem with the pseudo–RNG analysis, we reject his conclusions simply because they hold in favor of influence even in Monte Carlo–constructed unperturbed data.

Other Published Comments on DAT

We have found five other published papers that either criticize or test DAT. Two experimental AP studies tested DAT when the targets systems were biological (Braud and Schlitz, 1989, and Braud, 1990); both supported an influence hypothesis. Bierman (1988) attempted an experimental test, but the data did not show evidence of any anomaly. Walker (1987) claimed that the timing parameters in RNG and pseudo–RNG studies exceeded human physiological capabilities and therefore precluded DAT. Similarly, Vassy's pseudo–RNG study (1990) suggested that timing arguments also could not be accounted for by DAT. The timing arguments we presented above in our analysis of the Radin and May pseudo–RNG study (1986) negates Waker and Vassy's criticisms. The details of our rebuttals can be found in May, Spottiswoode, Utts, and James (1995).

Circumstantial Evidence Against an AP Model for RNG Data

Experiments with hardware RNG devices are not new. In fact, the title of Schmidt's very first paper on the topic (1969) portended our conclusion, "Precognition of a Quantum Process." Schmidt lists PK as a third option after two possible sources for precognition, and remarks, "The experiments done so far do not permit a distinction (if such a distinction is at all meaningful) between the three possibilities." From 1969 onward, the RNG research has been strongly oriented toward a PK model. The term micro–PK, itself, embeds the force concept further into the lexicon of RNG descriptions.

In this section, we examine a number of RNG experimental results that provide circumstantial evidence against the AP hypothesis. Any single piece of evidence could be easily dismissed; however, taken together, they demonstrate a substantial case against AP.

Internal Complexity of RNG Devices and Source Independence

Schmidt (1974) conducted the first experiment to explore potential dependencies upon the internal workings of his generators. Since by definition AP implies a force or influence, it seemed reasonable to expect that an influence should depend upon the details of the target system. In this study, one generator produced individual binary bits, which were derived from the β–decay of ^{90}Sr, while the other "binary" output was a majority vote from 100 bits, each of which were derived from a fast electronic diode. Schmidt reports individually significant effects with both generators, yet does not observe a significant difference between the generators.

This particular study is interesting, quite aside from the timing and majority vote issues; the binary streams were derived from fundamentally different physical sources. Radioactive β–decay is governed by the weak nuclear force, and electronic devices (e.g., noise diodes) are governed by the electromagnetic force. Schematically speaking, the electromagnetic force is approximately 1,000 times as strong as the weak nuclear force, and modern high–energy physics has shown them to be fundamentally different after about 10^{-10} seconds after the big bang (Raby, 1985). Thus, a putative AP–force would have to interact equally with these two forces; and since there is no mechanism known that will cause the electromagnetic and weak forces to interact with each other, it is unlikely that AP will turn out to be the first coupling mechanism. The lack of difference between β–decay and noise diode generators was confirmed years later by May, Humphrey, and Hubbard (1980).

We have already commented upon one aspect of the timing issue with regard to Radin and May's (1986) experiment and the papers by Walker (1987) and Vassy (1990). May (1975) introduced a scheme to remove any first–order biases in binary generators that also is relevant to the timing issue. The output of his generator was a match or anti–match between the random bit stream and a target bit. One mode of the operation of the device, which May describes, included an oscillating target bit — one oscillation per bit at approximately 1 MHz rate.[5] May and Honorton (1975) and Honorton and May (1975) reported significant effects with the RNG operating in this mode. Thus, significant effects can be seen even with devices that operate in the microsecond time domain, which is three orders of magnitude faster than any known physiological process.

Effects with Pseudorandom Number Generators

Pseudorandom number generators are, by definition, those that depend upon an algorithm, which is usually implemented on a computer. Radin (1985) analyzed all the pseudo–RNGs commonly in use and found that they require a starting value (i.e., a seed), which is often derived from the comput-

[5]Later, this technique was adopted by Jahn (1982).

er's system clock. As we noted above, Radin and May (1986) showed that the bit stream, which proved to be "successful" in a pseudo–RNG study, was bit–for–bit identical with the stream, which was generated later, but with the same seed. With that generator, at least, there was no change from the expected bit stream. Perhaps it is possible that the seed generator (i.e., system clock) was subjected to some AP interaction. We propose two arguments against this hypothesis:

(1) Even one cycle interruption of a computers' system clock will usually invoke a system crash; an event not often reported in pseudo–RNG experiments.

(2) Computers use crystal oscillators as the basis for their internal clocks. Crystal manufacturers usually quote errors in the stated oscillation frequency of the order of 0.001 percent. That translates to 500 cycles for a 50 MHz crystal, or to 10 ms in time. Assuming that the quoted error is a 1 a estimate, and that a putative AP interaction acts at within the $\pm 2\sigma$ domain, then shifting the clock by this amount might account for only one seed shift in Radin and May's experiment. By Monte Carlo methods, we determined that, given a random entry into seed–space, the average number of ticks to reach a "significant" seed is 10; therefore, even if AP could shift the oscillators by 2-0, it cannot account for the observed data.

Since computers in pseudo–RNG experiments are not reported as "crashing" often, it is safe to assume that pseudo–RNG results are only due to AC. In addition, since the results of pseudo–RNG studies are statistically inseparable from those reported with true RNGs, it is also reasonable to assume that the mechanisms are similarly AC–based.

Precognitive AC

Using the tools of modem meta–analysis, Honorton reviewed the precognition card–guessing database (Honorton and Ferarri, 1989). This analysis included 309 separate studies reported by 62 investigators. Nearly two million individual trials were contributed by more than 50,000 subjects. The combined effect size was $\varepsilon = 0.020 \pm 0.002$, which corresponds to an overall combined effect of 11.40. Two important results emerge from Honorton's analysis. First, it is often stated by critics that the best results are from studies with the least methodological controls. To check this hypothesis, Honorton devised an eight–point quality measure (e.g., automated recording of data, proper randomization techniques) and scored each study with regard to these measures. There was no significant correlation between study quality and study score. Second, if researchers improved their experiments over time, one would expect a significant correlation of study quality with date of publication. Honorton found $r = 0.246$, $df = 307$, $p = 2 \times 10^{-7}$. In brief, Honorton concludes that a statistical anomaly exists in this data that cannot be explained by poor study quality or a large variety of other hypotheses including the file drawer; therefore, a potential mechanism underlying DAT has been verified.

SRI International's RNG Experiment

May, Humphrey, and Hubbard (1980) conducted an extensive RNG study at SRI International in 1979. They applied state–of–the–art engineering and methodology to construct two true RNGs, one based on the β–decay of ^{137}Pm and the other based on an MD–20 noise diode from Texas Instruments. It is beyond the scope of this paper to describe, in detail, the intricacies of this experiment; however, we will discuss those aspects that are pertinent to this discussion.

Technical Details

Each of the two sources were battery operated and optically coupled to a Digital Equipment Corporation LSI 11/23 computer. Fail–safe circuitry would disable the sources if critical physical parameters (e.g., battery voltages and currents, temperature) exceed preset ranges. Both sources were subjected to environmental testing which included extreme temperature cycles, vibration tests, E&M and nuclear gamma and neutron radiation tests. Both sources behaved as expected, and the critical parameters, such as temperature, were monitored and their data stored along with the experimental data.

A source was sampled at 1 KHz rate. After eight milliseconds, the resulting byte was sent to the computer while the next byte was being obtained. In this way, a continuous stream of 1 ms data was presented to the computer. May et al. had specified, in advance, that bit number 4 was the designated target bit. Thus each byte provided 3 ms of bits prior to the target and 4 ms of bits after the target bit.

A trial was defined as a definitive outcome from a sequential analysis of bit four from each byte. In exchange for not specifying the number of samples in advance, sequential analysis requires that the Type I and Type II errors, and the chance and extra–chance hitting rate be specified in advance. In May et *al.*'s two–tailed analysis, $\alpha = \beta = 0.05$ and the chance and extra–chance hitting rate was 0.50 and 0.52, respectively. The expected number of samples to reach a definitive decision was approximately 3,000. The outcome from a single trial could be in favor of a hitting rate of 0.52, 0.48, or at chance of 0.50, with the usual risk of error in accordance with the specified Type I and Type II errors.

Each of seven operators participated in 100 trials of this type. For an operator's data to reach independently statistical significance, the operator had to produce 16 successes in 100 trials, where a success was defined as extra–chance hitting (i.e., the exact binomial probability of 16 successes for 100 trials with an event probability of 0.10 is 0.04 where one less success is not significant). Two operators produced 16 and 17 successful trials, respectively.

Temporal Analysis

We analyzed the 33 trials from the two independently significant operators from May ***et al.***'s experiment. Each of the **33** trials consisted of approximately 3,000 bits of data with –3 bits and +4 bits of 1 ms/bit temporal history surrounding the target bit. We argue that if the significance observed in the target bits was because of AP, we would expect a large correlation with the target bit's immediate neighbors, which are only ±1 ms away. As far as we know, there is no known physiological process that can be cognitively, or in any other way, manipulated within a millisecond. We might even expect a 100% correlation under the complete AP model.

We computed the linear correlation coefficients between bits 3 and 4, 4 and 5, and 3 and 5. For binary data:

$$N\phi^2 = \chi^2(df = 1),$$

where ϕ is the linear correlation coefficient and N is the number of samples. Since we examined three different correlations for 33 trials, we computed 99 different values of $N\phi^2$. Four of them produced χ^2 values that were significant —well within chance expectation. The complete distribution is shown in Figure 8. We see that there is excellent agreement of the 99 correlations with the χ^2 distribution for one degree of freedom, which is shown as a smooth curve.

We conclude, therefore, that there was no evidence beyond chance to suggest that the target bit neighbors were affected even when the target bit analysis produced significant evidence for an anomaly. So, knowing the physiological limitations of the human systems, we further concluded that the observed effects could not have arisen due to a human–mediated force (i.e., AP).

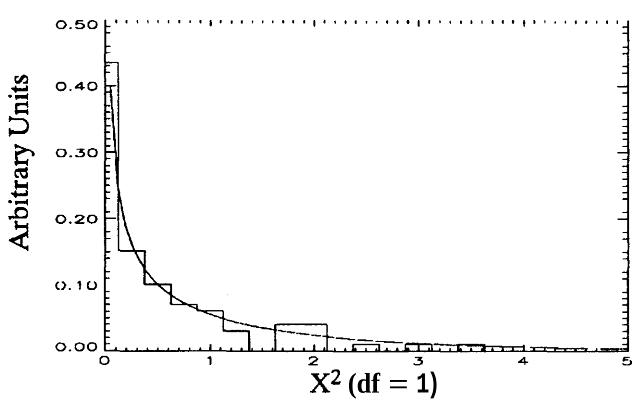

Fig. 8. Observed and Theoretical Correlation Distributions.

Mathematical Model of the Noise Diode

Because of the unique construction parameters of Texas Instrument's MD–20 noise diode, May ***et al.*** were able to construct a quantum mechanical model of the detailed workings of the device. This model contained all known properties of the material and it's construction parameters. For example, the band gap energy in Si, the effective mass of an electron or hole in the semiconductor, and the impurity concentration were among the parameters for the model. The model was successful at calculating the diode's known and measured behavior as a function of temperature. May ***et al.*** were able to simulate their RNG experiment down to the quantum mechanical details of the noise source. They hoped that by adjusting the model's parameters so that the computed output agreed with the experimental one, that they could gain insight as to where the influence "entered" the device.

May ***et al.*** were not able to find a set of model parameters that mimicked their RNG data. For example, changing the band gap energy in Si for short periods of time; increasing or reducing the electron's effect mass; or redistributing or changing the impurity content produced no unexpected changes in the device output. The only device behavior that could be effected was its known function of temperature.

Because of the construction details of the physical RNG, this result could have been anticipated. The changes that could be simulated in the model were all in the microsecond domain because of the details of the device. Both with the RNG and in its model, the diode's multi–MHz output was filtered by a 100–KHz wide bandwidth filter. Thus, any microsecond changes would not pass through the filter. In short, because of this filtering, the RNG was particularly insensitive to potential changes of the physical parameters of the diode.

Yet solid statistical evidence for an anomaly was seen by May ***et al.*** Since the diode device was shown mathematically and empirically to be insensitive to environmental and physical changes, these results must have been as a result of AC rather than AP. In fact, this observation coupled with the bit timing argument, which we have described above, led May et al. to question force–like models in RNG studies in general.

Summary of Circumstantial Evidence Against AP

We have identified six circumstantial arguments that, when taken together, provide increasingly difficult requirements that must be met by a putative AP force. In summary, the RNG database demonstrates that:

(1) Data are independent of internal complexity of the hardware RNG device.
(2) Data are independent of the physical mechanism producing the randomness (i.e., weak nuclear or electromagnetic).
(3) Effects with pseudorandom generators are statistically equivalent to those observed with true hardware generators.

(4) Reasonable AP models of mechanism do not fit the data.
(5) In one study, bits which are ±1 ms from a "perturbed" target bit are themselves unperturbed.
(6) A detailed model of a diode noise source, which includes all known physics of the device, could not simulate the observed data streams.

In addition, AC, which is a mechanism to describe the data, has been confirmed in non–RNG experiments. We conclude, therefore, an AP force that is consistent with the database must

- Be equally coupled to the electromagnetic and weak nuclear forces.
- Be mentally mediated in times shorter than one millisecond.
- Follow a $1/n^{1/2}$ law.

For these to be true, an AP force would be at odds with an extensive amount of verified physics and common behavioral observables. We are ***not*** saying, therefore, that it cannot exist; rather, we are suggesting that instead of having to force ourselves to invent a whole new science, we should look for ways in which AP might fit into the present world view. In addition we should invent information–based and testable alternate mechanisms for the experimental observable~.

Discussion

We now address the possible n–dependence of the model parameters. A degenerate case arises if ε_{AP} is proportional to $1/n^{1/2}$; if that were the case, we could not distinguish between the micro–AP model and DAT by means of tests on the n-dependence of results. If it were the case that in the analysis of the data from a variety of experiments, participants, and laboratories, the slope of a Z^2 versus n linear least–squares fit were zero, then either $\varepsilon_{AP} = 0.0$ or ε_{AP} is proportional to $1/n^{1/2}$, the accuracy depending upon the precision of the fit (i.e., errors on the zero slope). An attempt might be made to rescue the micro–AP hypothesis by explaining the $1/n^{1/2}$ dependence of ε_{AP} in the degenerate case as a fatigue or some other time dependence effect. That is, it might be hypothesized that anomalous perturbation abilities would decline as a function of n; however, it seems improbable that a human–based phenomenon would be so widely distributed and constant and give the $1/n^{1/2}$ dependency in differing protocols needed to imitate DAT. We prefer to resolve the degeneracy by wielding Occam's razor: if the only type of anomalous perturbation which fits the data is indistinguishable from AC, and given that we have ample demonstrations of AC by independent means in the laboratory, then we do not need to invent an additional phenomenon called anomalous perturbation. Except for this degeneracy, a zero slope for the fit allows us to reject all micro–AP models, regardless of their n–dependencies.

DAT is not limited to experiments that capture data from a dynamic system. As we indicated above, DAT may also be the mechanism in protocols which utilize quasi–static target systems. In a quasi–static target system, a random process occurs only when a run is initiated; a mechanical dice thrower is an example. Yet, in a series of unattended runs of such a device there is always a statistical variation in the mean of the dependent variable that may be due to a variety of factors, such as Brownian motion, temperature, humidity, and possibly the quantum mechanical uncertainty principle (Walker, 1974). Thus, the results obtained will ultimately depend upon when the run is initiated. It is also possible that a second–order DAT mechanism arises because of protocol selection; how and who determines the order in tri–polar protocols. In second order DAT there may be individuals, other than the formal subject, whose decisions effect the experimental outcome and are modified by AC. Given the limited possibilities in this case, we might expect less of an impact from DAT.

In surveying the range of anomalous mental phenomena, we reject the evidence for experimental macro–AP because of poor artifact control and accept the evidence for precognition and micro–AP because of the large number of studies and the positive results of the meta–analyses. We believe that DAT, therefore, might be a general model for anomalous mental phenomena in that it reduces mechanisms for laboratory phenomena to only one — the anomalous transtemporal acquisition of information.

Our recent results in the study of anomalous cognition (May, Spottiswoode, and James, 1994) suggest the the quality of AC is proportional to the change in Shannon entropy. Following Vassy (1990), we compute the change in Shannon entropy for an extra–chance, binary sequence of length *n*. The total change of entropy is given by:

$$\Delta S = S_0 - S$$

where for an unbiased binary sequence of length n, $S_0 = n$, and S is given by:

$$S = -np_1 \log_2 p_1 - n(1-p_1)\log_2(1-p_1).$$

Let $p_1 = 0.5\,(1 + \varepsilon)$ and assume that ε, the effect size, is small compared to one (i.e., typical RNG effect sizes are of the order of 3×10^{-4}). Using the approximation:

$$\ln(1+\varepsilon) = \varepsilon - \frac{\varepsilon^2}{2},$$

we find that S is given by:

$$S = n - n\frac{\varepsilon^2}{2\ln 2},$$

or that the total change of entropy for a biased binary sequence is given by;

$$\Delta S = S_o - S = n\frac{\varepsilon^2}{2\ln 2}$$

Since our analysis of the historical RNG database shows that the effect size is proportional to $1/n^{1/2}$, the total change of Shannon entropy becomes a constant that is independent of the sequence length:

$$\Delta S = \text{constant}$$

We have seen in our other AC experiments that the quality of the data is proportional to the change of the target entropy. In RNG experiments the quality of the data is equivalent to the excess hitting, which according to DAT is mediated by AC and should be independent of the sequence length. We have shown above that the quality of RNG data depends upon the change of target entropy and is independent of the sequence length. Therefore we suggest that the change of target entropy may account for successful AC and RNG experiments.

Conclusions

When DAT is applied to the RNG database, a simple force–like perturbational model fails, by many orders of magnitude, as a viable candidate for the mechanism. In addition, when viewed along with the collective circumstantial arguments against a force–like explanation, it is clear that another model is required. Any new model must explain why quadrupling the number of bits in the sequence length fails to produce a Z–score twice as large.

Given that one possible information mechanism (i.e., precognitive AC) can, and has been, independently confirmed in the laboratory, and given the weight of the empirical, yet circumstantial, arguments taken together against AP, we conclude that the anomalous results from the RNG studies arise not because of a mentally mediated force, but rather because of a human ability to be a mental opportunist by making AC–mediated decisions to capitalize on the locally deviant circumstances.

Generally, we suggest that future RNG and pseudo–RNG studies be designed in such a way that the criteria, as outlined in this paper and in May, Utts, Spottiswoode (1995), conform to a valid DAT analysis. Parapsychology has evolved to the point where we can no longer be satisfied with yet one more piece of evidence of a statistical anomaly. We must identify the sources of variance as suggested by May, Spottiswoode, and James (1994); limit them as much as possible; and apply models, such as DAT, which can begin to shed light on the physical, physiological, and psychological mechanisms of anomalous mental phenomena.

Acknowledgements

Since 1979, there have been many individuals who have contributed to the development of DAT. We would first like to thank David Saunders without whose remark this work would not have been. Beverly Humphrey kept the philosophical integrity intact at times under extreme duress. We are greatly appreciative of Zoltán Vassy, to whom we owe the Z–score formalism, to George Hansen, Donald McCarthy, and Scott Hubbard for their constructive criticisms and support. Finally, we acknowledge the contribution of two anonymous referees. In our opinion, the review process works. Our revised version of this paper is vastly improved over the original submission.

References

Bierman, D. J. (1988). Testing the IDS model with a gifted subject. ***Theoretical Parapsychology,*** 6, 31.

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

Braud, W. G. and Schlitz, M. J. (1989). Possible role of intuitive data sorting in electrodermal biological psychokinesis (Bio–PK). ***The Journal of the American Society for Psychical Research,*** 83, 4, 289.

Braud, W. G. (1990). Distant mental influence of rate of hemolysis of human red blood cells. ***The' Journal of the American Society for Psychical Research,*** 84, 1, 1.

DeGroot, Morris H. (1985). ***Probability and Statistics, 2nd Edition.*** Reading, MA: Addison–Wesley Publishing Co.

Dobyns, Y. H. (1993). Selection versus influence in remote REG anomalies. ***Journal of Scientific Exploration,*** 7, 3, 259.

Honorton, C. and May, E. C. (1975). Volitional control in a psychokinetic task with auditory and visual feedback. ***Research in Parapsychology,*** 1975, 90.

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

Honorton, C., Berger, R. E., Varvoglis, M. P., Quant, M., Derr, P., Schechter, E. I., and Ferrari, D. C. (1990). Psi Communication in the Ganzfeld. ***Joumal of Parapsychology,*** 54, 99.

Hubbard, G. S., Bentley, P. P., Pasturel, P. K., and Isaacs, J. (1987). A remote action experiment with a piezoelectric transducer. ***Final Report - Objective H, Task 3 and 3a, Project 1291.*** Menlo Park, CA: SRI International.

Hyman, R. and Honorton, C. (1986). A joint communiqué: The psi ganzfeld controversy. ***Journal of Parapsychology,*** 50, 351.

Jahn, R. G. (1982). The persistent paradox of psychic phenomena: an engineering perspecitve. ***Proceedings of the IEEE,*** 70, 2, 136.

Jahn R. G. and Dunne, B. J. (1986). On the quantum mechanics of consciousness, with application to anomalous phenomena. ***Foundations of Physics,*** 16, 8, 721.

Lewis, T. G. (1975). ***Distribution Sampling for Computer Simulation.*** Lexington, MA: Lexington Books.

May, E. C. (1975). PSIFI: A physiology–coupled, noise–driven random generator to extend PK studies. ***Research in Parapsychology,*** 1975, 20.

May, E. C. and Honorton, C. (1975). A dynamic PK experiment with Ingo Swann. ***Research in Parapsychology,*** 1975, 88.

May, E. C., Humphrey, B. S., Hubbard, G. S. (1980). Electronic system perturbation techniques. ***Final Report.*** Menlo Park, CA: SRI International

May, E. C., Radin, D. I., Hubbard, G. S., Humphrey, B. S., and Utts, J. (1985). Psi experiments with random number generators: an informational model. ***Proceedings of Presented Papers Vol I.*** The Parapsychological Association 28th Annual Convention, Tufts University, Medford, MA, 237.

May, E. C., J. M. Utts, and S. J. Spottiswoode (1995). Decision augmentation theory: Toward a model of anomalous mental phenomena. Accepted for publication in the ***Journal of Parapsychology.***

May, E. C., S. J. Spottiswoode, J. M. Utts, and C. L. James (1995). Applications of decision augmentation aheory. Accepted for publication in the ***Journal of Parapsychology.***

May, E. C., Spottiswoode, S. J., and James, C. L. (1994). Shannon entropy as a possible intrinsic target property. ***The Journal of Parapsychology,*** 58,384.

Nelson, R. D., Jahn, R. G., and Dunne, B. J. **(1986).** Operator–related anomalies in physical systems and information processes. *Journal of the Society for Psychical Research,* **53,803,261.**

Nelson, R. D., Dobyns, Y. H., Dunne, B. J., and Jahn, R. G., and **(1991).** Analysis of variance of REG experiments: Operator intention, secondary parameters, database structures. *PEAR* Laboratory *Technical Report 91004,* School of Engineering, Princeton University.

Raby, S. **(1985).** Supersymmetry and cosmology. in *Supersymmetry, Supergravity, and Related Topics.* Proceedings of the XVth GIFT International Seminar on Theoretical Physics, Sant Feliu de Guixols, Girona, Spain. World Scientific Publishing Co. Pte. Ltd. Singapore, **226.**

Radin, D. I. **(1985).** Pseudorandom number generators in Psi research. *Journal of Parapsychology,* 49, 4, 303.

Radin, *D.* I. and May, E. C. **(1986).** Testing the intuitive data sorting mode with pseudorandom number generators: A proposed method. *The Proceedings of Presented Papers of the 29th Annual Convention of the Parapsychological Association,* Rohnert Park, CA: Sonoma State University.

Radin, D. I. and Nelson, R. D. **(1989).** Evidence for consciousness–related anomalies in random physical systems. *Foundations of Physics,* **19, 12, 1499.**

Rosenthal, R. **(1991).** Meta–analysis procedures for social research. *Applied Social Research Methods Series, Vol. 6.* Newbury Park, CA: Sage Publications.

Schmidt. H. **(1969).** Precognition of a quantum process. *Journal of Parapsychology,* 33, 2, 99.

Schmidt. *H.* **(1974).** Comparison of PK action on two different random number generators. *Journal of Parapsychology,* **38,** 1, 47.

Stanford, R. G. (1974a). An experimentally testable model for spontaneous psi events I. Extrasensory events. *Journal of the American Society for Physical Research,* 68, 34.

Stanford, R. G. (1974b). An experimentally testable model for spontaneous psi events II. Psychokinetic events. *Journal of the American Society for Physical Research,* **68,321.**

Stanford, R. G., Zenhausern R., Taylor, A., and Dwyer, M. A. **(1975).** Psychokinesis as psi–mediated instrumental response. *Journal of the American Society for Physical Research,* **69, 127.**

Stokes, D. M. **(1987).** Theoretical parapsychology. In *Advances in Parapsychological Research 5.* Jefferson, NC: McFarland & Company, Inc.

Utts, J. **(1991).** Replication and meta–analysis in parapsychology. *Statistical Science,* **6**, 4, **363.**

Vassy, **Z. (1990).** Experimental study of precognitive timing: Indications of a radically noncausal operation. *Journal of Parapsychology,* **54,299.**

Walker, E. H. **(1974).** Foundations of Paraphysical and Parapsychological phenomena. *Proceedings of an International Conference: Quantum Physics and Parapsychology.* Oteri, E. Ed. New York, NY: Parapsychology Foundation, Inc.

Walker, E. H. **(1987).** A comparison of the intuitive data sorting and quantum mechanical observer theories. *Journal of Parapsychology,* **51,217.**

Walker, E. **H. (1984).** A review of criticisms of the quantum mechanical theory of psi phenomena. *Journal of Parapsychology,* **48,277.**

Washburn S. and Webb, R. A. **(1986).** Effects of dissipation and temperature on macroscopic quantum tunneling in Josephson junctions. In *New Techniques and Ideas in Quantum Measurement Theory.* Greenburger, *D. M.* Ed. New York, NY: New York Academy of Sciences, **66.**

Appendix

Mathematical Derivations for the Decision Augmentation Theory

In this appendix, we develop the formalism for the Decision Augmentation Theory (DAT). We consider cases for mean chance expectation, force-like interactions, and informational processes under two assumptions — normality and Bernoulli sampling. For each of these three models, we compute the expected values of $\mathbf{Z}$ and $\mathbf{Z}^2$, and the variance of $\mathbf{Z}^2$.[6]

[6] We wish to thank Zoltán Vassy for originally suggesting the Z^2 formalism.

Mean Chance Expectation

Normal Distribution. We begin by considering a random variable, X, whose probability density function is normal, (i.e., $N(\mu_0, \sigma_0^2)$).[7] After many unbiased measures from this distribution, it is possible to obtain reasonable approximations to μ_0 and σ_0^2 in the usual way. Suppose n unbiased measures are used to compute a new variable, Y, given by:

$$Y_k = \frac{1}{n}\sum_{j=1}^{n} X_{jk}.$$

Then Y is distributed as $N(\mu_0, \sigma_n^2)$, where $\sigma_n^2 = \sigma_0^2/n$. If Z is defined as

$$Z = \frac{Y_k - \mu_0}{\sigma_n},$$

then Z is distributed as $N(0,1)$ and $E(Z)$ is given by:

$$E_{MCE}^{N}(Z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z e^{-0.5z^2} dz = 0.$$

Since $Var(Z) = 1 = E(Z^2) - E^2(Z)$, then

$$E_{MCE}^{N}(Z^2) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^2 e^{-0.5z^2} dz = 1.$$

The $Var(Z^2) = E(Z^4) - E^2(Z^2) = E(Z^4) - 1$. But

$$E_{MCE}^{N}(Z^4) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^4 e^{-0.5z^2} dz = 3.$$

So

$$Var_{MCE}^{N}(Z^2) = 2.$$

Bernoulli Sampling. Let the probability of observing a one under Bernoulli sampling be given by p,. After n samples, the discrete Z-score is given by:

$$Z = \frac{k - np_0}{\sigma_0\sqrt{n}},$$

where

$$\sigma_0 = \sqrt{p_0(1-p_0)},$$

[7]Throughout **this appendix, this notation means:**

$$N(\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-0.5\left[\frac{x-u}{\sigma}\right]^2\right)$$

and k is the number of observed ones ($0 \leq k \leq n$). The expected value of Z is given by:

$$E^B_{MCE}(Z) = \frac{1}{\sigma_0\sqrt{n}}\sum_{k=0}^{n}(k - np_0)B_k(n, p_0), \tag{1}$$

where

$$B_k(\text{n}, p_0) = \binom{n}{k} p_0^k (1 - p_0)^{n-k}.$$

The first term in Equation 1 is $E(k)$ which, for the binomial distribution, is np_0.

Thus

$$E^B_{MCE}(Z) = \frac{1}{\sigma_0\sqrt{n}}\sum_{k=0}^{n}(k - np_0)B_k(n, p_0) = 0.$$

The expected value of Z^2 is given by:

$$\begin{aligned} E^B_{MCE}(Z^2) &= Var(Z) + E^2(Z), \\ &= \frac{Var(k - np_0)}{n\sigma_0^2} + 0, \\ &= \frac{n\sigma_0^2}{n\sigma_0^2} = 1. \end{aligned}$$

As in the normal case, the $Var(Z^2) = E(Z^4) - E^2(Z^2) = E(Z^4) - 1$. But[8]

$$E^B_{MCE}(Z^4) = \frac{1}{n^2\sigma_0^4}\sum_{k=0}^{n}(k - np_0)^4 B_k(n, p_0)$$

$$E^B_{MCE}(Z^4) = 3 + \frac{1}{n\sigma_0^2}(1 - 6\sigma_0^2)$$

So,

$$Var^B_{MCE}(Z^2) = 2 + \frac{1}{n\sigma_0^2}(1 - 6\sigma_0^2) = 2\left(1 - \frac{1}{n}\right), \quad (p_0 = 0.5).$$

Force-Like Interactions

Normal Distribution. Under the perturbation assumption described in the text, we let the mean of the perturbed distribution be given by $\mu_0 + \varepsilon_{AP}\sigma_0$, where

[8] **Johnson, N. L. and Kotz, S., (1969).** ***Discrete Distributions.*** **John Wiley & Sons, New York, p. 51.**

may be a function of n and time. The parent distribution for the random variable, X, becomes $N(\mu_0+\varepsilon_{AP}\sigma_0, \sigma_0^2)$. As in the mean-chance-expectation case, the average of n independent values of X is $Y \sim N(\mu_0+\varepsilon_{AP}\sigma_0, \sigma_n^2)$. Let

$$y = \mu_0 + \varepsilon_{AP}\sigma_0 + \Delta y.$$

For a mean of n samples, the Z-score is given by

$$Z = \frac{y-\mu_0}{\sigma_n} = \frac{\varepsilon_{AP}\sigma_0 + \Delta y}{\sigma_n} = \varepsilon_{AP}\sqrt{n} + \zeta,$$

where ζ is distributed as $N(0,1)$ and is given by $\Delta y/\sigma_n$. Then the expected value of Z is given by

$$E_{AP}^{N}(Z) = E\left(\varepsilon_{AP}\sqrt{n} + \zeta\right) = \varepsilon_{AP}\sqrt{n} + E(\zeta) = \varepsilon_{AP}\sqrt{n}\ ,$$

and the expected value of Z^2 is given by

$$E_{AP}^{N}(Z^2) = E\left(\left[\varepsilon_{AP}\sqrt{n} + \zeta\right]^2\right) = n\varepsilon_{AP}^2 + E(\zeta^2) + 2\varepsilon_{AP}\sqrt{n}\,E(\zeta) = 1 + \varepsilon_{AP}^2 n,$$

since $E(\zeta) = 0$ and $E(\zeta^2) = 1$.

In general, Z^2 is distributed as a non-central χ^2 with one degree of freedom and non-centrality parameter $n\varepsilon_{AP}^2$. Thus, the variance of Z^2 is given by[9]

$$Var_{AP}^{N}(Z^2) = 2\left(1 + 2n\varepsilon_{AP}^2\right).$$

Bernoulli Sampling. As before, let the probability of observing a one under mean chance expectation be given by p_0 and the discrete Z-score be given by:

$$Z = \frac{k - np_0}{\sigma_0\sqrt{n}},$$

where k is the number of observed ones ($0 \le k \le n$). Under the perturbation assumption, we let the mean of the distribution of the single-bit probability be given by $p_1=p_0+\varepsilon_{AP}\sigma_0$, where ε_{AP} is an anomalous-perturbation strength parameter. The expected value of Z is given by:

$$E_{AP}^{B}(Z) = \frac{1}{\sigma_0\sqrt{n}}\sum_{k=0}^{n}(k - np_0)B_k(n, p_1)\ ,$$

where

[9]Johnson, N. L. and Kotz, S,. (1970). ***Univariate Distributions — 2.*** **John Wiley & Sons, New York. p. 134.**

$$B_k(n, p_1) = \binom{n}{k} p_1^k (1 - p_1)^{n-k}.$$

The expected value of Z becomes

$$E_{AP}^B(Z) = \frac{1}{\sigma_0 \sqrt{n}} \left[\sum_{k=0}^{n} k B_k(n, p_1) - n p_0 \right]$$

$$= \frac{(p_1 - p_0)\sqrt{n}}{\sigma_0} = \varepsilon_{AP} \sqrt{n}.$$

Since in general $Z = \varepsilon\, n^{1/2}$, ε_{AP} can be considered an AP effect size. The expected value of Z^2 is given by:

$$E_{AP}^B(Z^2) = Var(Z) + E^2(Z),$$

$$= \frac{Var(k - np_0)}{n\sigma_0^2} + \varepsilon_{AP}^2 n,$$

$$= \frac{p_1(1 - p_1)}{\sigma_0^2} + \varepsilon_{AP}^2 n.$$

Expanding in terms of $p_1 = p_0 + \varepsilon_{AP}\sigma_0$,

$$E_{AP}^B(Z^2) = 1 + \varepsilon_{AP}^2 (n - 1) + \frac{\varepsilon_{AP}}{\sigma_0}(1 - 2p_0)$$

If $p_0 = 0.5$ (i.e., a binary case) and $n \gg 1$, then the expectation value for the binomial case reduces to the same as in the normal case.

We begin the calculation of $Var(Z^2)$ by using the equation for the jth moment of a binomial distribution

$$m_j = \frac{d^j}{dt^j}\left[\left(q + pe^t\right)^n \right] \Big|_{t=0}$$

Because $Var(Z^2) = E(Z^4) - E^2(Z^2)$, we must evaluate $E(Z^4)$. Or

$$E_{AP}^B(Z^4) = \frac{1}{n^2 \sigma_0^4} \sum_{k=0}^{n} (k - n p_0)^4 B_k(n, p_1).$$

Expanding

$$\frac{(k - n p_0)^4}{n^2 \sigma_0^4},$$

using the appropriate moments, and subtracting $E^2(Z^2)$, yields

$$Var(Z^2) = C_0 + C_1 n + C_{-1} n^{-1}.$$

Where

$$C_0 = 2 - 36\varepsilon_{AP}^2 + 10\varepsilon_{AP}^4 + 8\frac{\varepsilon_{AP}}{\sigma_0}(1-2p_0)(1-2\varepsilon_{AP}^2) + 6\left[\frac{\varepsilon_{AP}}{\sigma_0}\right]^2$$

$$C_1 = 4\varepsilon_{AP}^2(1-\varepsilon_{AP}^2) + 4\frac{\varepsilon_{AP}^3}{\sigma_0}(1-2p_0),$$

and

$$C_{-1} = 48 - 6[\varepsilon_{AP}^2 - 3]^2 + 12\frac{\varepsilon_{AP}^3}{\sigma_0}(1-2p_0) + \frac{(1-7\varepsilon_{AP}^2)}{\sigma_0^2} + \frac{\varepsilon_{AP}}{\sigma_0^3}(1-2p_0)(12p_0^2 - 12p_0 + 1).$$

Under the conditions that $\varepsilon_{AP} \ll 1$ (a frequent occurrence in many experiments), and if we ignore any terms of higher order than $\varepsilon_{AP}{}^2$, then the variance reduces to

$$Var(Z^2) = 2 - 36\varepsilon_{AP}^2 + 8\frac{\varepsilon_{AP}}{\sigma_0}(1-2p_0) + 6\left[\frac{\varepsilon_{AP}}{\sigma_0}\right]^2 + 4\varepsilon_{AP}^2 n +$$
$$\frac{1}{n}\left[-6 + 36\varepsilon_{AP}^2 + \frac{(1-7\varepsilon_{AP}^2)}{\sigma_0^2} + \frac{\varepsilon_{AP}}{\sigma_0^3}(1-2p_0)(12p_0^2 - 12p_0 + 1)\right].$$

We notice that when $\varepsilon_{AP} = 0$, the variance reduces to the **mean-chance-expectation** case for Bernoulli sampling. When $n \gg 1$, $\varepsilon_{AP} \ll 1$, and $p_0 = 0.5$, the variance reduced to that derived under the normal distribution assumption. Or,

$$Var_{AP}^B(Z^2) \approx 2(1 + 2n\varepsilon_{AP}^2).$$

Informational Process

Normal Distribution. The primary assumption in this case is that the parent distribution remains unchanged, (i.e., $(N(\mu_0,\sigma_0{}^2))$. We further assume that because of an anomalous-cognition-mediated bias, the sampling distribution is distorted leading to a Z-distribution of $N(\mu_z,\sigma_z{}^2)$. In the most general case, μ_z and σ_z may be functions of *n* and time.

The expected value of *Z* is given by definition as

$$E_{AC}^N(Z) = \mu_z.$$

The expected value of Z^2 is given by definition as

$$E_{AC}^N(Z^2) = \mu_z^2 + \sigma_z^2.$$

The $Var(Z^2)$ can be calculated by noticing that

$$\frac{Z^2}{\sigma_0^2} \sim \mathrm{X}_{nc}^2\left(1, \frac{\mu_z^2}{\sigma_z^2}\right).$$

So the $Var(Z^2)$ is given by

$$Var\left(\frac{Z^2}{\sigma_0^2}\right) = 2\left(1 + 2\frac{\mu_z^2}{\sigma_z^2}\right)$$

Or

$$Var_{AC}^N(Z^2) = 2\left(\sigma_z^4 + 2\mu_z^2\sigma_z^2\right).$$

Bernoulli Sampling. As in the normal case, the primary assumption is that the parent distribution remains unchanged, and that because of an AC-mediated bias the sampling distribution is distorted leading to a discrete Z-distribution characterized by μ_z and σ_z. Thus, by definition, the expected values of Z and Z^2 are given by

$$E_{AC}^B(Z) = \mu_z ,$$

and

$$E_{AC}^B(Z^2) = \mu_z^2 + \sigma_z^2 ,$$

respectively. For any value of n, estimates of these parameters are calculated from N data points as

$$\mu_z = \frac{1}{N}\sum_{j=1}^{N} z_j ,$$

and

$$\sigma_z^2 \approx \frac{N}{N-1}\left[\sum_{j=1}^{N}\frac{z_j^2}{N} - \left(\frac{1}{N}\sum_{j=1}^{N} z_j\right)^2\right].$$

The $Var(Z^2)$ for the discrete case is identical to the continuous case. Therefore

$$Var_{AC}^B(Z^2) = 2\left(\sigma_z^4 + 2\mu_z^2\sigma_z^2\right).$$