

The MegaREG Experiment: Replication and Interpretation

Y. H. DOBYNS, B. J. DUNNE, R. G. JAHN, AND R. D. NELSON

*Princeton Engineering Anomalies Research
School of Engineering and Applied Science
Princeton University
Princeton, NJ 08544-5263*

Abstract—Anomalous effects of human intention on the output of electronic random event generators (REGs) have been well established at the PEAR laboratory and elsewhere. A simple model of this effect as a change in the binary probability of the REG digits would predict that larger statistical yield can be achieved simply by processing more bits. This hypothesis was explored previously using protocols ranging from 20 to 2000 bits per trial, with results that were consistent with the bitwise model, but had too little resolution to rule out many competing models. More recently, a “MegaREG” experiment was deployed to test this hypothesis using 2-million-bit trials interspersed with 200-bit trials in a double-blind protocol.

In the initial phase of MegaREG, the 200-bit trials produced outcomes comparable with our standard experiments, while the 2-million-bit trials produced an effect somewhat larger in absolute scale, but inverted with regard to intention. A subsequent replication phase reproduced these findings, except for statistically nonsignificant quantitative changes. These appear to be secondary consequences of a statistically significant difference between operators having, and lacking, prior experience in REG experiments, the relative proportions of which account for the differences between these experimental phases. Other operator population distinctions, such as gender, and various secondary protocol parameters, had no significant effects.

A related experiment called “MegaMega,” differing from MegaREG only in that all data used 2 million bits per trial, with no interspersal of a second data type, produced a reversed intentional effect of the same scale. It also displayed a significant asymmetry between the intentional runs and the non-intentional baselines, which was not seen in MegaREG.

The combined result of all high-speed experiments was an effect size per trial of -2.77 ± 0.69 times that seen in earlier REG experiments, but given the larger number of bits per trial, the bitwise probability change was some 30 times smaller. The composite score for the intentional effect in high-density data across all experiments was $T = -4.03$ (d.f. $> 10^5$), $p = 5.65 \times 10^{-5}$ (2-tailed). The causes of the change of scale, and of the inversion of sign in the effect, remain unknown. Explanations that can be ruled out with a high degree of confidence include statistical artifact, the change in the source, the use of different operator pools, and the double-blind interspersal of data types. Testable explanations that remain potentially viable include increased task complexity, inherent timing or rate limits on anomalous functioning, and changes in the psychological environment.

Keywords: consciousness—anomalies—human/machine—interaction—random event—generators—replicability

Introduction

The Princeton Engineering Anomalies Research (PEAR) laboratory has conducted extensive experimental and theoretical study of anomalous effects of human consciousness on various types of random event generators (REGs) since 1979 (Jahn *et al.*, 1997). This work builds on previous and ongoing studies by many other researchers, and is particularly close in design and protocol to that of Schmidt (1970a,b). An extensive meta-analysis by Radin and Nelson (1989) found that most of these experimental programs, including PEAR's, produced anomalous effects of broadly similar nature and scale.

The anomalous effect seen in these experiments consists of a shift in the mean output level of REGs that correlated with pre-stated human intention. Thus, regardless of mechanism or model, any database containing an anomaly displays a change in the empirical probability of the individual binary events comprising that database. Most of the PEAR data are consistent with the hypothesis that the anomalous effect is in fact nothing more nor less than an alteration of the probability of elementary binary events in the experiment, rather than some more complicated process which would produce the empirical probability shift as a consequence (Dobyys, 2000; Jahn *et al.*, 1991). For reasons of protocol standardization, however, many physical and psychological variables were held constant in these experiments, which means that the consistency with the probability-change hypothesis could be an artifact arising from the uniformity of some other key parameter throughout the experiments. If the anomalous effect is truly a shift in the elementary bit-level probability, however, the statistical yield of an experiment should be increased by increasing the number of bits processed, while holding all other factors constant.

This prospect was explored in a preliminary fashion using various protocols. The standard "REG200" protocol collects the sum of 200 random bits into a single "trial," at a sampling rate of 1000 bits/second. Thus, data collection is active for 0.2 second, a period of time easily perceptible to the operator. An intervening pause of approximately 0.7 second leads to a mean data generation rate of approximately 0.9 second/trial. Two exploratory variants, labeled REG20 and REG2000, collected respectively 1/10 and 10 times as many bits per trial from the same noise source, with the sampling rate changed correspondingly so that the periods of trial accumulation were the same. Although the amount of data accumulated in these protocols was small relative to the primary REG200 experiment, the results seemed compatible with a bitwise effect hypothesis (Dobyys, 2000; Jahn *et al.*, 1997). These previous explorations are summarized in Figure 1, showing all three data points well within a 95% confidence interval

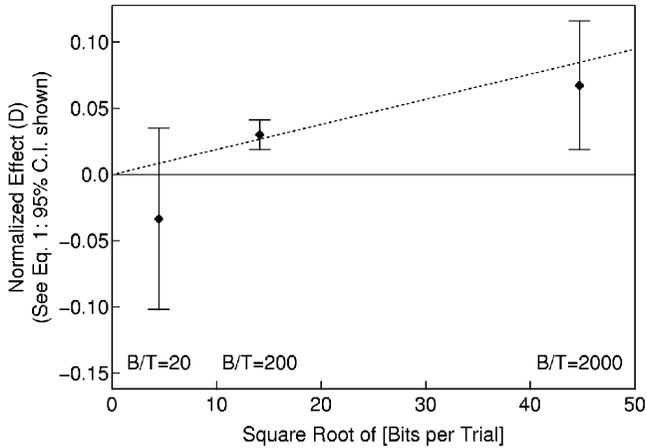


Fig. 1. Original REG investigations of bit rate effects.

(1.96 times as wide as the 1σ error bars drawn) of the sloping dotted line indicating constant effect per bit.

The original “MegaREG” experiment (Ibison, 1997) attempted a more thorough test of the bitwise effect hypothesis. The concept was to increase the standard trial size from a sum of 200 bits to a sum of 2,000,000 bits. If the effect was a direct alteration of binary probability, then increasing the number of bits per experiment by 10^4 , while keeping all other parameters of the experiment constant, should increase the statistical yield one hundredfold. Several alternative models of the nature of the anomalous REG effect were considered as well, each of which predicted a different effect size for this change in bit density.

In order to accomplish this major increase in the number of bits processed while still presenting the trials to the operator at the same rate as in earlier experiments, a high-speed noise source unique to this experiment was developed, as described in detail in Ibison (1998). To ensure that the subjective participation of operators was changed as little as possible, the user interface for the data-collection program was unaltered from earlier REG experiments. As before, trials were collected and presented to the operator at a pace of about one per second. In any given experimental series, the operator had the option of (a) seeing the numerical trial value, (b) seeing a graphical cumulative-deviation trace, or (c) seeing no feedback at all until the end of the run.

To control for possible psychological differences in the operators’ mental state induced by the awareness of a different experiment and/or by feedback with numerically large trial values (mean value 1,000,000 instead of 100), MegaREG was designed to generate both the new, 2-million-bit trials, and 200-bit trials comparable to those of earlier experiments. These were presented in indistinguishable formats, as detailed below, in a randomized order, to prevent any possibility of non-anomalous awareness by the operators of which trial type

currently was being processed. This was a “double-blind” randomization in that the experimenters, no less than the operators, remained uninformed about the source of each trial until the minimum designed database size had been accumulated. The unblinding of the experimenters by the analysis of the initial data defined the endpoint of the original experiment.

The same noise source was used for both types of trials. In each, two million bits were collected and processed. In the “high-density” trials, all bits were summed to create the raw trial value, a random integer with a theoretical mean of 1,000,000 and standard deviation 707.1. In the “low-density” trials, only every ten-thousandth bit of the set of 2 million was used for the sum. (The “low-density” nomenclature was chosen because of the fact that the source still ran at its full sampling rate, but only a few of the samples were actually used as data. However, the use of every ten-thousandth sample had the effect of temporally spacing all of the samples actually used as though they had been taken at a 1-kHz sampling rate, so that the “low-density” mode just as validly could be considered a “low-speed” mode.) The resulting low-density trial value was a random integer with theoretical mean 100 and standard deviation 7.071. Low-density and high-density trials were interspersed, in a pseudorandom pattern, in each “run” of the experiment as defined below.

To maintain the indistinguishability of the two trial types, a “normalized” value for the high-density trials was computed. If t was the original trial value, the normalized value was given by $0.01 \times (t - 1000000) + 100$; that is, the mean value was subtracted out, the resulting difference divided by 100 to reproduce the standard deviation of the low-density data, the low-density mean of 100 was added back, and the result rounded to the nearest integer. Thus, under the null hypothesis of no effect, the normalized value of a high-density trial should have had exactly the same distribution as a low-density trial: mean 100, standard deviation 7.071.

In all other respects, the MegaREG experiment followed the same protocol as other contemporaneous REG experiments: each “series” comprised 1000 trials in each of the three intentional conditions: high, baseline, and low. Trials were collected in “runs” of continuous data generation, which, according to the operator’s preference at the start of the series, could be either 1000 trials, in which case the whole series would comprise just one run in each intention, or 100 trials, in which case the series would require 30 runs, 10 in each intention. The pseudorandom lookup table that governed the interspersal of high and low densities guaranteed exactly 500 trials of each type in a 1000-trial run; it did not, however, maintain this exact balance in individual 100-trial runs. Hence the counts of individual high- and low-density trials in this experiment was not exactly divisible by 500 in each intention.

In the formulae and data tables to follow, the high-density data are presented and analyzed as the normalized trial values, rather than as raw data values, for two reasons. First, with both types of trial value reduced to the same range, the same algorithms and tests can be applied to both. Second, the normalization process has

no impact on the statistical yield of a given anomalous mean shift, and this statistical yield is the primary variable of interest (see Appendix). Since the statistical yield per trial is identical whether one uses the original raw trial values or their normalized forms, the convenience of using the normalized values becomes the deciding argument.

Besides the bitwise probability model for the REG anomaly, other models with different scaling properties were also deployed as possible predictions for the MegaREG outcome. They are not recounted in detail because they all have been rendered moot, *i.e.*, the initial experimental result refuted all of the proposed theories (including the null hypothesis of no effect). As reported in Ibison (1997, 1998), the low-density data produced an effect comparable in magnitude to that of earlier REG experiments, though not statistically significant due to the relatively small database. In contrast, the high-density data departed from chance expectation by more than 3σ , but in a direction *contrary* to the stated intention of the operators. Such a change of sign cannot be represented as a scaling effect, regardless of the model.

After the initial unblinding, data collection was left open for any operators who wished to generate data, and additional data generated in 1996 after the unblinding were included in the analyses and presentations of Ibison (1997, 1998). In hopes of understanding these curious results, a substantial replication database was generated in 1998 and 1999. The data from these three phases of data collection (original experiment as designed, post-unblinding period, and formal replication period) are the basis of the following discussion.

Analysis Variables

The high-speed data source used for both high- and low-density data is somewhat less stable than the older REG sources, and its output departs from the theoretical binomial distribution for trials. Hence, the actual source statistics must be determined empirically, so parametric statistics must be computed as Student's T -scores, based on the empirical standard deviation. Degrees of freedom (d.f.) are not reported explicitly for most results, since these are in the range 10^4 to 10^5 ; in this regime the T -distribution differs negligibly from the standard normal distribution, and p -values and confidence intervals can be computed from the latter without appreciable inaccuracy.

As a further safeguard against the lack of a sound theoretical distribution for the MegaREG source, we depart from PEAR's traditional presentation and analysis of the data in terms of the independent outcomes of the three intentions. Instead, we use two independent measures, which can be constructed from the three intentional datasets, and are provably immune to most of the artifacts that might arise from an unknown source distribution. Although there are many ways to construct two such independent measures from three raw data sources, their formulae are completely determined by adding two further constraints: each measure should have expectation 0 under the null hypothesis of no intentional

effect, and each should have the same standard deviation or standard error as the measurements from which it is computed.

Let h , l , and b denote one experimental outcome in each of the three intentions high, low, and baseline, respectively. Let each of these outcomes have a statistical observation uncertainty σ_h , σ_l , σ_b , respectively. Then the two measures,

$$\mathcal{D} \equiv \frac{h - l}{\sqrt{2}}, \quad \mathcal{A} \equiv \frac{h + l - 2b}{\sqrt{6}} \quad (1)$$

each have expectation 0, and their variances are

$$\sigma^2(\mathcal{D}) = \frac{\sigma_h^2 + \sigma_l^2}{2}; \quad \sigma^2(\mathcal{A}) = \frac{\sigma_h^2 + \sigma_l^2 + 4\sigma_b^2}{6}. \quad (2)$$

The symbols for these measures are chosen mnemonically. \mathcal{D} is the delta-effect, the difference between the oppositely directed active intentional conditions. \mathcal{A} is the asymmetry, the difference between the passive baseline and the two active intentions.

In addition to having well-defined distributions regardless of possible source irregularities, \mathcal{D} and \mathcal{A} are mutually independent, or at least uncorrelated. The proofs of these useful properties are given in the Appendix, along with empirical data regarding the validity of the mathematical assumptions used, and the combination formulae for calculating \mathcal{D} and \mathcal{A} properly (*i.e.*, in a manner immune to secular drift) in composite databases.

Experimental Results

[A series-by-series report of the raw data is available in the Appendices to Dobyys *et al.* (2002).]

As mentioned, there is some ambiguity in the exact boundary between the original and replication databases. Originally, the unblinding date (2 August 1996) was taken as the close of the initial experiment. Prior to unblinding, the identity of specific trials as high- or low-density had been as unknown to the experimenters as to the operators, and this was considered an important aspect of the experimental protocol. Nevertheless, the experiment was left open to operators who wished to contribute data, and three such volunteers generated six more series after the unblinding. These were included in the database reported by Ibison in 1997 and 1998. Therefore, on the basis of prior publication, these late data are part of the original experiment, whereas on the basis of protocol, they are part of the replication, conducted after the experimenters had discovered the peculiar contrast between high-density and low-density data.

Since the amount of data involved is small compared to those databases that are unambiguously part of the replication or of the original experiment, this issue of definitions might seem pedantic. Unfortunately, the late-1996 data contain such large effects that they have a disproportionate influence on any subset in which

TABLE 1
MegaREG Experimental Phases

Density	Delta-effect (\mathcal{D})	$T(\mathcal{D})$	Asymmetry	$T(\mathcal{A})$
Original experiment: 59 series				
Low	0.0341 ± 0.0404	0.8456	-0.0443 ± 0.0404	-1.0948
High	-0.1200 ± 0.0403	-2.9802	-0.0195 ± 0.0401	-0.4854
Post-unblinding data: 6 series				
Low	0.2265 ± 0.1251	1.8100	0.0841 ± 0.1269	0.6628
High	-0.3048 ± 0.1254	-2.4302	0.0733 ± 0.1250	0.5861
Replication experiment: 84 series				
Low	-0.0056 ± 0.0329	-0.1713	0.0295 ± 0.0329	0.8950
High	-0.0679 ± 0.0327	-2.0779	0.0375 ± 0.0327	1.1464
Combined results: 149 series				
Low	0.0189 ± 0.0250	0.7546	0.0033 ± 0.0250	0.1335
High	-0.0971 ± 0.0249	-3.9043	0.0170 ± 0.0248	0.6865
Difference T -scores: original–replication				
$T_d(\text{Low})$	0.7638		-1.4141	
$T_d(\text{High})$	-1.0033		-1.1007	
$T_d(\Delta\mathcal{D})$	1.2492		NA	
Original–post-unblinding				
$T_d(\text{Low})$	-1.4628		-0.9641	
$T_d(\text{High})$	1.4033		-0.7063	
$T_{\mathcal{D}}(\Delta\mathcal{D})$	-2.0266		NA	
Post-unblinding–replication				
$T_d(\text{Low})$	1.7941		0.4167	
$T_d(\text{High})$	-1.8276		0.2769	
$T_{\mathcal{D}}(\Delta\mathcal{D})$	2.5609		NA	

they are included. The choice of including them in the replication or in the original database thus has a substantial impact on the statistical relationship between the two. Given this fact, and the arguments above for distinguishing the late-1996 data both from the original experiment and from the later replication, the responsible course is to report the total experiment in *three* subdivisions: the data generated prior to the 2 August 1996 unblinding; the post-unblinding data generated later in 1996; and the formal replication effort of 1998 and 1999.

Table 1 presents the results of the MegaREG experiment in all three phases. Figure 2 displays the overall \mathcal{D} and \mathcal{A} values, with associated uncertainties, for the high- and low-density data in each of the three phases. In Figure 2, \mathcal{D} is plotted on the horizontal axis (labeled “Differential Effect”), while \mathcal{A} is plotted on the vertical axis (labeled “Asymmetry”). Error bars (1σ) are presented along both axes. The error bars on the low-density data are marked with arrowheads, while the error bars on the high-density data are marked with terminal crossbars.

Table 1 includes only 59 series in the original MegaREG database, whereas the previous publications on this database (Ibison, 1997, 1998) list 70 series. The

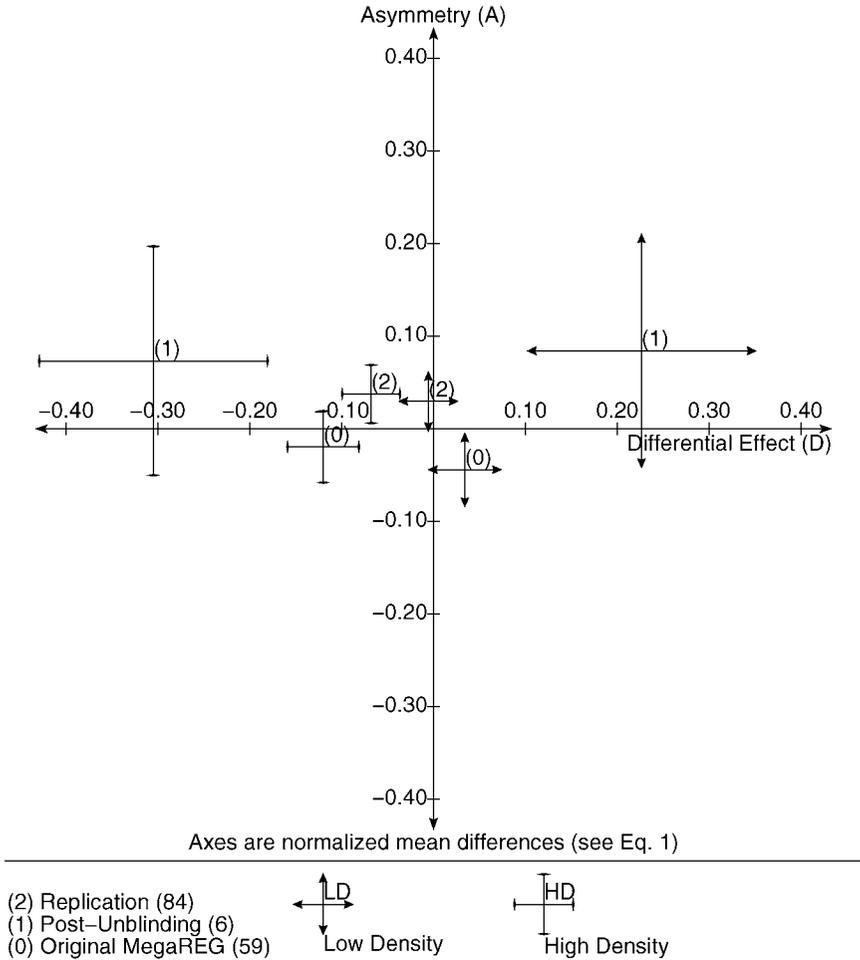


Fig. 2. MegaREG experimental phases.

discrepancy has two sources. One is the separation of the six post-unblinding series from the original database. In addition, five series run in late 1995 as pilot tests, before the hardware design was finalized, were included in the earlier analysis although they properly should not be considered part of the formal database. The removal of these 11 series does not produce any qualitative change in the results of the original database.

The difference *T*-scores in the last section of Table 1 are calculated according to the formula:

$$T_d = \frac{m_1 - m_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \tag{3}$$

where m_1 and m_2 denote two measured values with observation uncertainties σ_1 and σ_2 . In general we do not have a directed hypothesis for T_d , so its sign is irrelevant but any p -values obtained for it must be two-tailed.

In interpreting Table 1, we first note that the asymmetry parameter \mathcal{A} seems to show only chance behavior. None of its direct T -scores are significant, in either density, nor are there significant differences between experimental phases. In contrast, for \mathcal{D} , each of the three phases has a negative value in the high density that is independently significant by a two-tailed, $p = 0.05$ criterion ($|T| > 1.96$). The low-density data show nonsignificant positive results in the original experiment and the post-unblinding phase, and a nonsignificant negative result in the replication.

The difference T -scores in the lower part of Table 1 show no $p < 0.05$ difference for either data density between any two experiment phases. However, this section adds another parameter. The difference between low-density Δ and high-density Δ , in any given phase, produces an observable quantity which we may call $\Delta\mathcal{D}$, expressing the change in performance between the two densities. The $\Delta\mathcal{D}$ rows in the last section then give the difference scores T_d for $\Delta\mathcal{D}$ between the indicated experimental phases. Note that the quantity on lines labeled $T_d(\Delta\mathcal{D})$ is thus a third-order difference on the raw data: \mathcal{D} is intrinsically a difference comparing high and low intentions; $\Delta\mathcal{D}$ is the difference between \mathcal{D} in the high and low data densities; and $T_d(\Delta\mathcal{D})$ is a difference between $\Delta\mathcal{D}$ in two phases of the experiment.

While $T_d(\Delta\mathcal{D})$ between the original and replication data is well within the range of normal chance variation, the post-unblinding phase does differ significantly from both of the other phases, at $p = 0.043$ for the original and $p = 0.010$ for the replication (both p -values 2-tailed). Thus, while neither the high-density nor the low-density conditions differ individually between any two of the three experimental phases, the post-unblinding phase becomes statistically distinguishable from both the original and the replication phases when the split between low- and high-density data is examined. This is due to the combination of increased \mathcal{D} in low-density and decreased \mathcal{D} in high-density, though neither change is significant by itself. The original and replication phases, on the other hand, remain statistically indistinguishable, even in the $\Delta\mathcal{D}$ parameter.

Since the original and replication data are statistically indistinguishable on every measure, it is safe to pool them for the subsidiary analyses to be conducted later. The status of the post-unblinding subset is more problematic. It is distinguishable from the rest of the data on one parameter, $\Delta\mathcal{D}$, but only on that parameter. It is a very small database, and 5 of its 6 series were generated by operators who did not participate in either of the other phases of the experiment. Although one might argue that the disparate post-unblinding subset should be discarded as an outlier, it seems safer to include it in the experimental data, at least provisionally. The fact that it is dominated by operators not appearing elsewhere in the experiment suggests that its peculiarities may be driven by idiosyncratic properties of those operators. If so, exclusion of these data would lead to inaccurate interpretations of the range of operator performances.

As the “Combined Results” in Table 1 illustrate, when all data are pooled, the counter-intentional result in the high-density \mathcal{D} achieves a T -score of -3.9043 ($p = 9.4 \times 10^{-5}$, 2-tailed).

Subsidiary Analyses

Aside from the “bottom line” results presented in Table 1 and Figure 2, we may hope for some illumination into the nature of the phenomenon to emerge from more detailed analysis. As with most PEAR experiments, a number of secondary parameters were examined along with the primary intentional variable. In addition, prior experimental experience strongly suggests that different operators may produce different results, whether as idiosyncratic individuals or as members of subgroups with shared properties (Dunne, 1991, 1998; Jahn *et al.*, 1991, 1997; Nelson *et al.*, 2000). In pursuing the following analyses, we use data pooled across all experimental phases.

Individual Operators

[Full data on individual operator performances can be found in the Appendices of Dobyys *et al.* (2002).]

Figure 3 presents a scatter plot showing all of the individual operator performances. Low-density data are presented with open circles, high-density with filled circles; to improve visibility, the error bars have been omitted. Note that the scale of Figure 3 has been expanded, compared to Figure 2, to accommodate the wide dispersion of individual results. (The appearance of only 23 low-density datasets is due to an overlap: the open circle for one operator’s low-density data [$\mathcal{D} = -0.0042$, $\mathcal{A} = 0.0678$] is invisible behind another operator’s high-density data [$\mathcal{D} = -0.0032$, $\mathcal{A} = 0.0676$].)

In the low-density data, the individual operator performances show a slight, nonsignificant bias toward positive \mathcal{D} , with 14 of 24 having $\mathcal{D} > 0$. In the high-density data, 20 of the 24 operators show $\mathcal{D} < 0$; this population imbalance itself is improbable with $p = 0.003$, 2-tailed. (A further Bonferroni correction for the fact that such an imbalance might have appeared in either density condition yields $p = 0.006$, still highly significant.) The fact that most operators have negative \mathcal{D} suggests that the reversal of intentional effect in the MegaREG database is a broadly distributed phenomenon among the operator population, rather than being driven by a few exceptional operators.

The wide scatter of the individual operator databases in Figure 3, relative to the overall summaries of Figure 2, is due in part to the smaller size, and hence reduced statistical resolution, of the individual databases. It is important to know, however, whether the actual amount of scatter is statistically distinguishable from that expected for normal random variation given the database sizes. The amount of variation among operators can be calculated by computing a χ^2 value for the departure of the various operators from the collective mean. If there are N

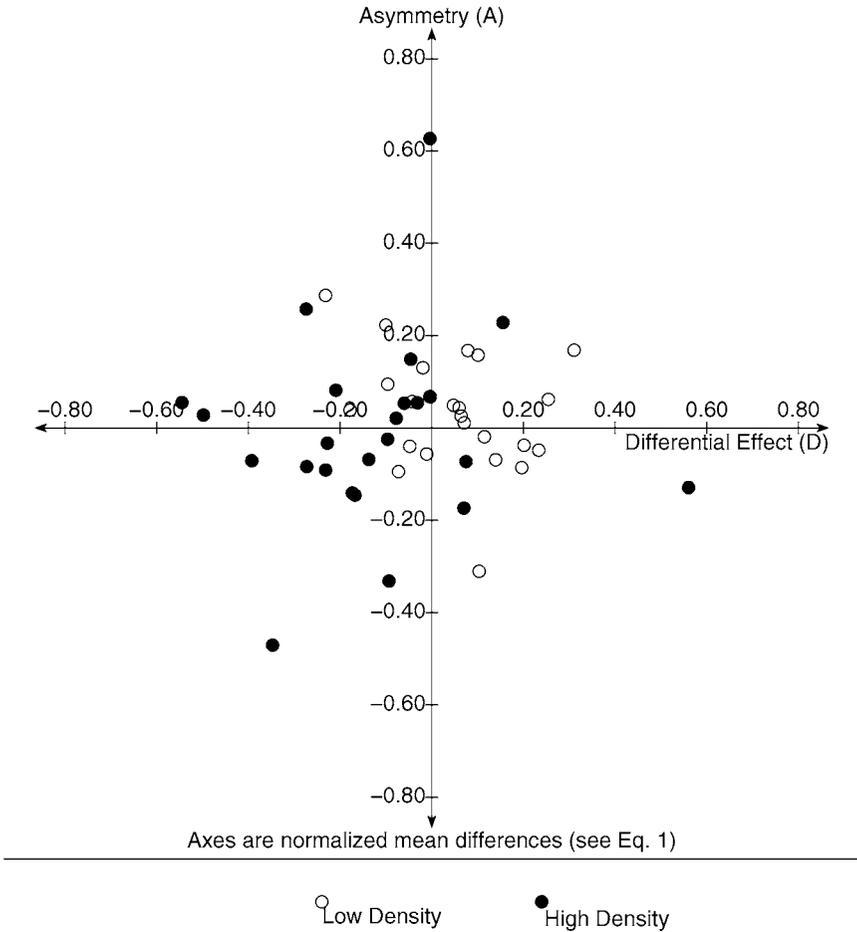


Fig. 3. Individual operator performances in MegaREG.

operators, and the i th operator's database has mean m_i with standard error σ_i , and if the composite mean for all operators is m , then

$$\chi^2 = \sum_{i=1}^N \left(\frac{m_i - m}{\sigma_i} \right)^2 \tag{4}$$

is distributed as a χ^2 with $N - 1$ d.f., under the null hypothesis that all operators have an identical effect on the data and the differences between operators are due to random variation.

Table 2 gives the inter-operator χ^2 values as computed from Eq. 4. The p -values given are upper-tail, describing the null-hypothesis likelihood of a larger χ^2 . Since we would, however, find suppressed variation equally as interesting as

TABLE 2
Inter-Operator Variability

Density	d.f.	$\chi^2(\mathcal{D})$	p	$\chi^2(\mathcal{A})$	p
Low	23	16.897	0.814	14.504	0.912
High	23	27.885	0.220	26.717	0.268

excess variation, for a 5% criterion we should look for either $p < .025$ or $p > .975$. It is evident that none such appear. Thus, any idiosyncratic individual variation among operators is too small for detection within the available statistical resolution.

Operator Populations

Rather than varying individually and idiosyncratically, operator performances may be divisible into two (or more) distinct populations. Because of the larger databases, a difference between two populations may be more readily detectable when the data from operators in each population are pooled, even though the individual-operator χ^2 measures do not show detectable increases in the overall dispersion of operator performances. One operator-distinguishing parameter that has proven to be important in many other experiments is operator gender (Dunne, 1998). Examining gender forces us to examine operator plurality as well. Some “operators” are actually pairs of operators working together (Dunne, 1991). This co-operator subgroup is treated here as a third “gender,” since all pairs consisted of one male and one female operator.

Figure 4 illustrates the results of dividing the MegaREG data into the three gender-based sub-populations: females, males, and co-operators. There are evident differences among the three populations, although the error bars make it clear that these may not be significant. Most striking is the fact that the inverted intentional effect in the high-density condition is almost twice as large for males as for females, and twice as large again for the co-operators. (This last result, though not statistically robust due to the large uncertainties, is consistent with earlier findings regarding co-operator pairs [Dunne, 1991].)

Another division that may have greater explanatory power was intended to test the possibility that the differences between MegaREG and earlier REG experiments were due to the differences in the operator pool. Specifically, the MegaREG operator population was subdivided into those operators with previous REG experience, and those operators whose first experience of REG-class experiments was MegaREG. Figure 5 displays \mathcal{D} and \mathcal{A} for these two populations of “previously experienced” and “new” operators, with the original REG200 results shown as a comparison benchmark. We may note in passing that the negative result in high-density data cannot be explained entirely by the change in operator pool, since both sets of high-density data remain individually

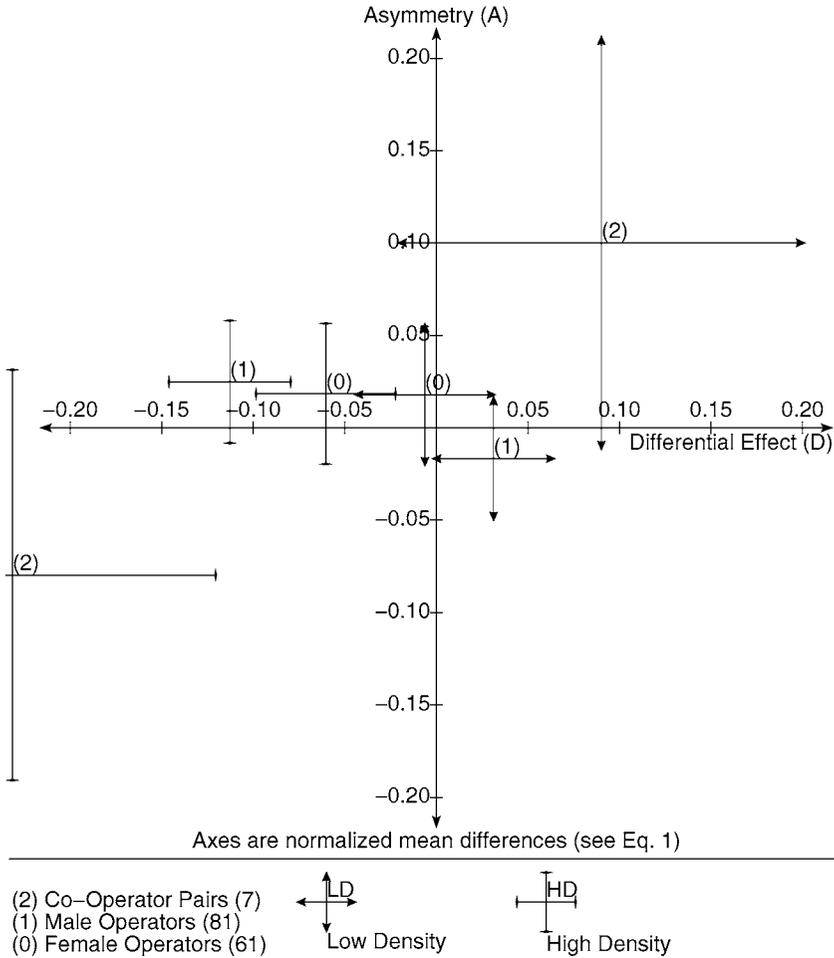


Fig. 4. Data from operator gender/number subgroups.

incompatible with the REG200 data. Of considerably greater interest, though, is the tremendous difference in performance between the experienced operators and the new operators.

It is evident in both Figures 4 and 5 that, just as in the whole-experiment summaries, \mathcal{A} is showing no detectable deviation from chance behavior. The statistical status of \mathcal{D} in these figures is not so obvious from inspection, so Table 3 presents the \mathcal{D} values for Figures 4 and 5 numerically. In addition to the value of \mathcal{D} in low- and high-density, the $\Delta\mathcal{D}$ parameter for the split between densities is reported. Table 3 also gives a χ^2 value for the differences among the three operator populations segregated by gender and number, and a T_d value for the two populations segregated by experience. The χ^2 values, which have 2 d.f.,

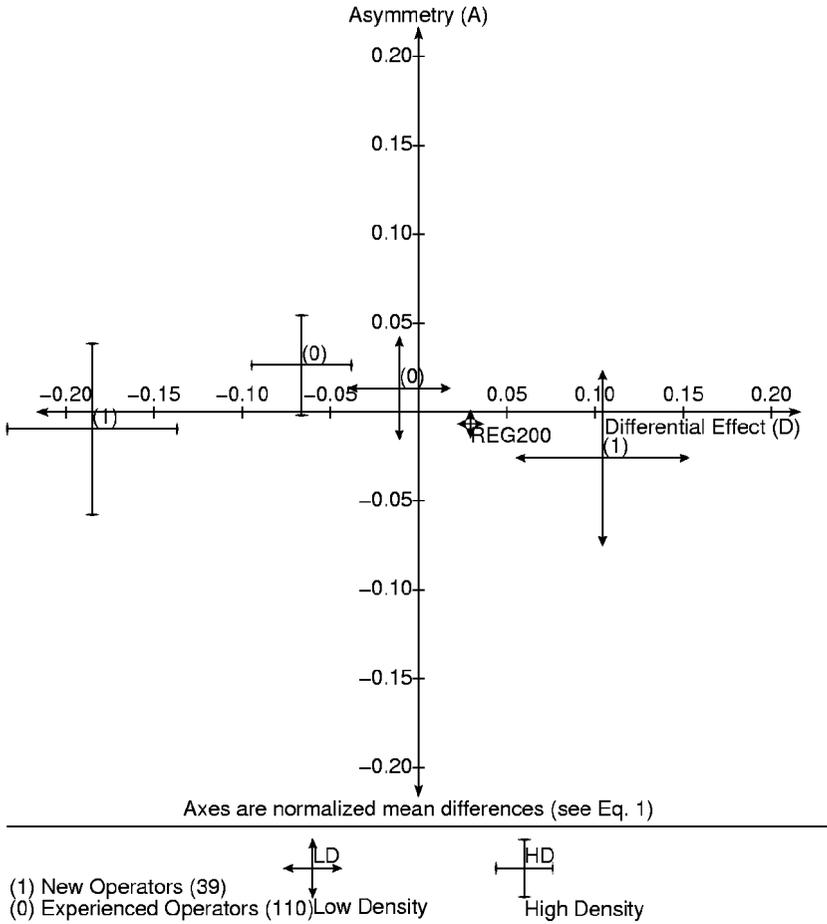


Fig. 5. New and experienced operators, with REG200 comparison.

TABLE 3
Operator Sub-Populations

Operator type	N_{ops}	N_{ser}	Low-density \mathcal{D}	High-density \mathcal{D}	$\Delta\mathcal{D}$
Females	7	61	-0.0062 ± 0.0391	-0.0604 ± 0.0388	0.0541 ± 0.0550
Males	12	81	0.0313 ± 0.0340	-0.1128 ± 0.0339	0.1441 ± 0.0480
Co-op pairs	5	7	0.0901 ± 0.1116	-0.2315 ± 0.1116	0.3216 ± 0.1578
Difference $\chi^2(p)$			0.9542 (0.621)	2.5626 (0.278)	3034 (0.192)
Experienced	11	110	-0.0109 ± 0.0290	-0.0664 ± 0.0289	0.0555 ± 0.0409
New	13	39	0.1044 ± 0.0492	-0.1850 ± 0.0489	0.2894 ± 0.0693
$T_d(p)$, Exp - New			$-2.0182 (0.043)$	$2.0987 (0.036)$	$-2.9045 (0.0037)$

TABLE 4
Individuals within Populations

Operator type	N_{ops}	$N_+ N_-$, low-density	$N_+ N_-$, high-density
All	24	14 10	4 20
Female	7	4 3	1 6
Male	12	8 4	2 10
Co-op	5	2 3	1 4
Experienced	11	5 6	2 9
New	13	9 4	2 11

indicate that despite the apparent scale of the differences between operator genders, their statistical significance cannot be confirmed. (Neither do any of the individual T_d values between pairs of operator types attain a $p < 0.05$ discrimination.) In contrast, the difference between experienced and new operators is independently significant in each data density at a two-tailed $p < 0.05$ level, and achieves $p = 0.0037$ (2-tailed) in the $\Delta\mathcal{D}$ condition. Despite the impressive difference between the two populations, Table 3 also confirms that the negative value of high-density \mathcal{D} is a statistically robust feature of both; while the new operators have almost triple the negative effect of the experienced operators ($\mathcal{D} = -0.1850$ vs. $\mathcal{D} = -0.0664$), the latter's performance is still independently significant with $T = (-0.0664/0.0289) = -2.2975$ ($p = .022$).

The results of Table 3 suggest that the apparent differences among experimental phases seen in Table 1 and Figure 2 are, in all likelihood, driven mostly by the operator-experience variable. In the original experiment, the 59 series were almost evenly contributed by both operator types: 31 series were generated by experienced operators, 28 by new operators. In the post-unblinding period, with its much larger effects, 5 of the 6 series were generated by new operators. In the replication phase, with its smaller effects and much smaller $\Delta\mathcal{D}$, 78 of the 84 series were generated by experienced operators, and only six by new operators.

In the discussion of Figure 3, we noted that individual operator performances showed a statistically significant preference for $\mathcal{D} < 0$ in high-density data. Table 4 examines this individual-operator measure within the subpopulations of Table 3. The columns labeled $N_+ | N_-$ give the number of operators with positive and negative \mathcal{D} , respectively, in the high- and low-density data. The most salient feature of this dissection is the consistent tendency for individuals to have negative \mathcal{D} across all categories of gender and experience in the high-density data.

Secondary Parameters

A variety of protocol parameters could be adjusted to suit the operator's preferences and comfort, and these were examined as secondary independent variables, in hope that their impact or lack thereof on the experimental results

might convey some insight into the nature of the anomaly. Four sets of secondary parameter variations were used in MegaREG. The assignment of intention to individual runs could be volitional (operator chooses whether the next run will be high, low, or baseline, within the constraints of an overall balanced design), or instructed (choice was made automatically by the program, using a pseudorandom procedure). The control of trial generation within runs could be automatic (trials were generated sequentially by the program until the run is complete) or manual (after each trial, the program waited for an operator keypress before generating the next). The feedback to the operator could be graphic (cumulative deviation trace drawn on the screen), digital (screen presented current trial value and running mean numerically), or nonexistent (screen presented only the number of trials generated, with no information concerning results). Finally, the runs themselves could be short (100 trials, so a series consisted of 10 runs in each intention), or long (1000 trials, so a series consisted of 1 run in each intention). The results for these data subdivisions are presented in Table 5. Once again only \mathcal{D} is presented; although neither variable shows statistically robust distinctions in any parameter, there are some suggestive trends approaching significance in \mathcal{D} , while \mathcal{A} shows no noteworthy activity.

For \mathcal{D} as presented in the table, there apparently is also no discernible sensitivity to the assignment mode. Volitional and instructed data are statistically indistinguishable in every aspect. In contrast, the negative high-density \mathcal{D} value is strikingly enhanced in the manual-control data, with an apparent effect size almost three times as large as the automatic-control data. This difference is not statistically significant, however, due to the small size of the manual database.

The third section of Table 5 suggests that the reversed intentional effect, and the difference between the two data densities, vanishes almost entirely in the digital-feedback data; is considerably stronger in the graphic-feedback mode that comprises the bulk of the database; and is strongest of all when feedback is entirely removed. However, the complete non-significance of the χ^2 measures for inter-dataset variability make it doubtful that these apparent variations comprise a genuine pattern.

The final part of Table 5 indicates that short runs show larger positive effects in low density, and larger negative effects in high density. Although these differences are slightly short of statistical significance, they are suggestive in that several earlier REG experiments have observed the same pattern of non-significant enhancement effects in shorter runs (Jahn *et al.*, 2000; Nelson *et al.*, 2000). (A meta-analytic Stouffer Z-score combining the current result with those presented in the two foregoing references achieves a significant value of 2.094.)

In summary, despite some suggestive variations, none of the secondary parameters show clearly resolved, statistically unambiguous effects on the data. Thus, the main indications of the subsidiary analyses are:

TABLE 5
Secondary Parameter Comparisons

Parameter	N_{ser}	Low-density \mathcal{D}	High-density \mathcal{D}	$\Delta\mathcal{D}$
Instructed	73	-0.0006 ± 0.0359	-0.0870 ± 0.0357	0.0864 ± 0.0506
Volitional	76	0.0372 ± 0.0348	-0.1067 ± 0.0347	0.1440 ± 0.0492
T_d and $p(T)$		-0.7569 (0.449)	0.3971 (0.691)	-0.8166 (0.414)
Automatic	138	0.0200 ± 0.0259	-0.0861 ± 0.0258	0.1060 ± 0.0366
Manual	11	0.0044 ± 0.0938	-0.2467 ± 0.0948	0.2511 ± 0.1334
T_d and $p(T)$		0.1594 (0.873)	1.6347 (0.102)	-1.0492 (0.294)
Graphic	110	0.0220 ± 0.0289	-0.1044 ± 0.0287	0.1264 ± 0.0408
Digital	27	-0.0199 ± 0.0595	-0.0457 ± 0.0593	0.0258 ± 0.0841
No-feedback	12	0.0771 ± 0.0903	-0.1443 ± 0.0905	0.2214 ± 0.1278
$\chi^2(p_\gamma)$		0.8516 (0.653)	1.0876 (0.581)	1.8969 (0.387)
Short runs	56	0.0384 ± 0.0413	-0.1549 ± 0.0412	0.1933 ± 0.0583
Long runs	93	0.0076 ± 0.0314	-0.0639 ± 0.0312	0.0715 ± 0.0443
T_d and $p(T)$		0.5925 (0.554)	-1.7626 (0.078)	1.6634 (0.096)

- The negative effect in high-density data is broadly distributed in the operator population, and cannot be attributed to a subset of peculiar or idiosyncratic operators.
- Neither is this effect attributable to the change in operator pool from earlier REG experiments, since it is present and statistically significant both in operators with previous REG experience, and in operators whose participation in MegaREG was their first exposure to this class of experiments.
- The effect is, however, much larger in the data contributed by new operators, to a statistically significant degree.
- While there are tentative indications that some secondary parameters may affect the experimental performance, any such effects are below the threshold for confident statistical detection in these data.

Parallel Results: The “MegaMega” Experiment

A separate experiment exploring high data rates was undertaken concurrently with the replication phase of MegaREG. This experiment was identical to MegaREG in all regards except one: all data were normalized high-density data. No low-density data were collected, so there was no double-blind interspersal of two data types. This experiment, dubbed “MegaMega,” thus directly addressed the impact of this double-blinding on the MegaREG result. The results for this experiment are summarized in Table 6 and Figure 6.

Figure 6 displays the MegaMega results, along with the MegaREG overall results in both densities and those of the original REG200 experiment, which corresponds to MegaREG low density. Table 6 reports the overall results of MegaMega, and the breakdowns by various subsets. Since the database is considerably smaller than the MegaREG database, comprising only 39 series, statistical uncertainties are appreciably larger. An immediately striking result is

TABLE 6
MegaMega Results

Data subset	N_{op}	N_{ser}	\mathcal{D}	$T(\mathcal{D})$	\mathcal{A}	$T(\mathcal{A})$
All data	11	39	-0.0510 ± 0.0345	-1.4779	0.0753 ± 0.0345	2.1809
Operator categories						
Experienced	7	29	-0.0287 ± 0.0401	-0.7161	0.0730 ± 0.0402	1.8185
New	4	10	-0.1146 ± 0.0678	-1.6910	0.0819 ± 0.0677	1.2092
Female	5	22	-0.0381 ± 0.0466	-0.8190	0.1300 ± 0.0465	2.7946
Male	6	17	-0.0669 ± 0.0515	-1.2979	0.0082 ± 0.0515	0.1592

the emergence of a significant overall asymmetry parameter \mathcal{A} in the total database. Since \mathcal{A} does not come close to statistical significance in any phase of MegaREG (Table 1), this must be classified as a departure from MegaREG behavior. The breakdown by operator types indicates that this asymmetry is driven entirely by female operators, and does not seem to be affected by previous experience.

Since the only difference between MegaMega and MegaREG is the double-blind interspersal of low-density data in the latter, this asymmetry would seem to be attributable to the simpler technical design of the former. It is also suggestive that the asymmetry appears to be driven by female operators, insofar as a tendency for asymmetric performance by females has been noted in other experiments (Dunne, 1998). If indeed this sort of asymmetry is a characteristic female pattern, one might speculate that the double-blind, interspersed data generation of MegaREG may have caused it to be suppressed.

In \mathcal{D} , the net effect for MegaMega (-0.0510 ± 0.0345) is quite close to that seen in the high-density data of the concurrent MegaREG replication database (-0.0679 ± 0.0327), and is not statistically distinguishable from the high-density \mathcal{D} for MegaREG as a whole (-0.0971 ± 0.0249 ; difference $T_d = 0.6863$). Moreover, as Table 6 shows, the pattern of negative \mathcal{D} by operator category is the same in MegaMega as in high-density MegaREG; *i.e.*, the effect is somewhat stronger for male operators than females, and is overwhelmingly driven by new operators as distinct from previously experienced operators.

In terms of \mathcal{D} , then, MegaMega appears to be a successful replication, in the sense that it shows the same effects and the same internal pattern of effects as the high-density data from MegaREG. The implications of this similarity will be addressed further in the Discussion. If we consider the indistinguishability of the results to justify pooling of the data, MegaMega brings the combined high-density database to an effect of $\mathcal{D} = -0.0813 \pm 0.0202$, $T = -4.0272$, $p = 5.65 \times 10^{-5}$, 2-tailed.

Relation to REG200

Figure 6 also displays the REG200 result for comparison with the

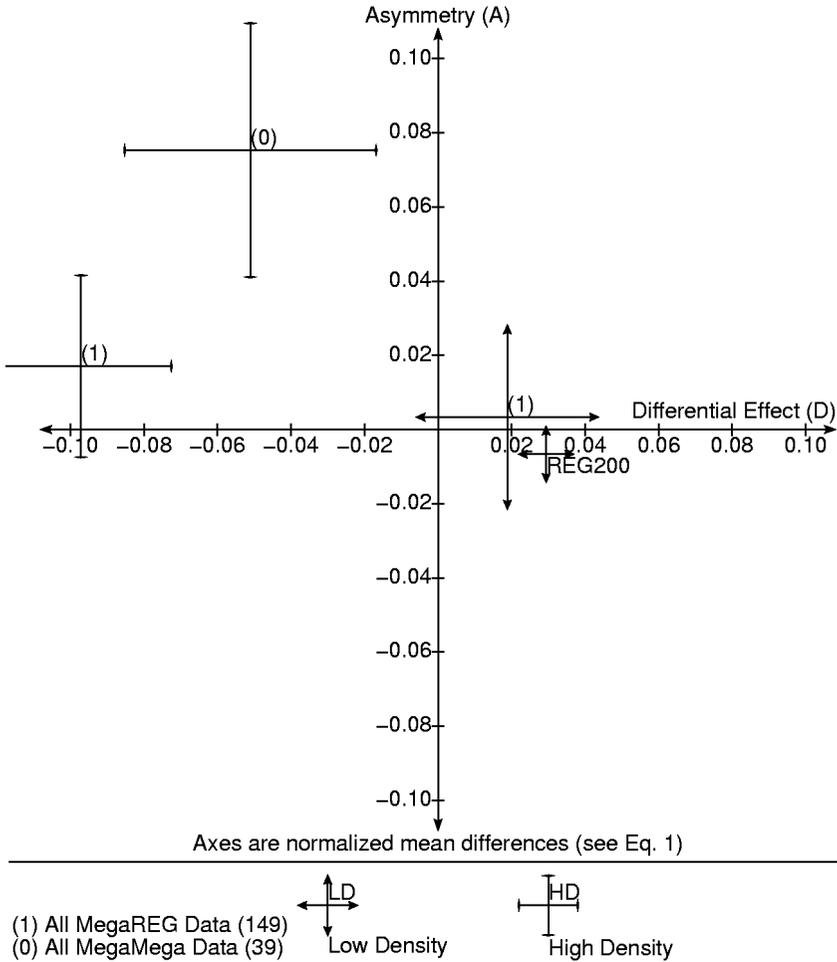


Fig. 6. MegaREG, MegaMega, and REG200.

MegaREG and MegaMega results, in the same \mathcal{D} , \mathcal{A} analysis variables. Like most MegaREG datasets, REG200 has an overall \mathcal{A} indistinguishable from 0 ($\mathcal{A} = -0.0067 \pm 0.0078$). Its intentional effect is $\mathcal{D} = 0.0294 \pm 0.0077$. This is statistically indistinguishable from the low-density \mathcal{D} , as is evident from the figure; using the value 0.0189 ± 0.0250 from Table 1, we obtain $T_q = 0.402$. The relatively large uncertainty in the low-density MegaREG result means, however, that while indistinguishable from REG200 it is also indistinguishable from zero; although in some sense the “least hypothesis” is that it should show the same effect size as the preceding experiment, this is not strongly favored over the hypothesis that it shows no effect at all.

The anti-intentional effect in high-density data is both highly significant

and clearly distinct from the intentional effect in REG200. The cause of the sign reversal remains a mystery, but the scale of the effect is clearly larger. The absolute value of the pooled MegaREG and MegaMega effect is $|\mathcal{D}| = 0.0813 \pm 0.0202$, some 2.77 ± 0.69 times as large as the REG200 effect, with $T_d = 2.401$ ($p = 0.016$, 2-tailed) for the difference in absolute magnitudes. The change in the bitwise effect can be extracted from the \mathcal{D} effect by recalling that the ten-thousandfold increase in bits per trial mandates a hundredfold increase in statistical yield per trial, if Δp is held constant. Therefore the effect per bit in high-density data is $(2.77/100) = 0.0277$ times the REG200 bitwise effect size, with $T_d = 3.711$ ($p = 2 \times 10^{-4}$, 2-tailed) for the difference between the two. The salient features of raising the bit count per trial from two hundred to two million thus seem to be (a) an inversion of the sign of the effect, (b) an approximately threefold increase in statistical yield per trial, and (c) an approximately thirtyfold decrease in statistical yield per bit, with all three relations being statistically well-established.

Discussion

Before the outcome of the original MegaREG experiment was known, the anticipatory theoretical efforts deployed for its interpretation involved various models for the expected scaling of the effect between low-density data, which corresponded to the vast majority of pre-existing PEAR REG data, and the high-density data with 10^4 times as many bits per trial. The simple model (Dobyns, 2000; Jahn *et al.*, 1997), which presumes that REG anomalies are due to a change in the probability of elementary binary events, predicts that the high-density data should have 100 times the statistical leverage of the low-density data. Various other models also were considered *a priori*, as detailed in Ibison (1997, 1998).

Unfortunately, the reversal of effect ($\mathcal{D} < 0$) in the high-density data refutes all of the proposed models, since none can accommodate a change of sign. On the other hand, the effect seen in the high-density data is sufficiently robust ($p = 5.65 \times 10^{-5}$) to render the null hypothesis untenable. All proposed models, including the null hypothesis, having been refuted, what then are the options for interpreting these data?

Some consideration not included in previous models must account for the reversal in sign. Ibison (1997, 1998) focuses theoretical interpretation on the “source-independence” theory promulgated by Schmidt, among others (Schmidt, 1974). It may indeed be the case that the difference between MegaREG and previous REG experiments resides in their noise sources. It is premature, however, to call this the only explanation available, or conclude that this experiment has addressed definitively the question of source independence. From “Combined Results” in Table 1 we see that the high-density data are not only statistically distinct from the chance expectation of 0, but also quite distinct from the low-density data generated on the same source: the difference is $T_d = 3.29$ ($p = 0.001$, 2-tailed). (Including the MegaMega data changes the

value slightly but not the conclusion: $T_d = 3.12$, $p = 0.002$.) If the $\mathcal{D} < 0$ effect in the high-density data is due to the different noise source, we must come to terms somehow with the fact that the *same* noise-generation hardware, using a different mode for postprocessing the raw data, has become “different” in terms of its anomalous performance. Another way to express the problem is that we find a theoretically unbiased sample of 1 bit in every 10,000 to be showing statistics completely incompatible with those of the parent noise stream from which the sample is drawn.

This difficulty with source-dependence can be avoided only by adopting a rather abstract notion of “source.” If one insists that the whole causal chain connecting the raw physical noise to the final recorded data must be regarded as the source, then the different postprocessing regimes qualify the high- and low-density modes of the MegaREG source as distinct “sources” in this extended sense of the term. While logically legitimate, this usage of “source” is somewhat counterintuitive. For example, one would normally consider the MegaREG “source” to be the separately powered, physically isolated unit which transmitted the raw bit-level signal to the experimental computer, but in contrast this alternative view includes as part of the “source” those operations, carried out entirely in software by the main computer, which distinguish a high-density trial from a low-density trial.

Setting aside source dependence, the following subsections outline some of the other explanations that have been considered seriously for the reversal of intentional effects in the high-density data.

Statistical Fluke

There was some initial concern that the odd result in the first MegaREG database somehow might be an artifact of inadequate data collection, despite the database having achieved its designed size. Moreover, the fact that the result was contrary to intention, while the initial models had presumed an intentional effect and had deployed one-tailed tests for its detection, made any statistical interpretation of the results problematic according to some schools of statistical inference. Both of these concerns have been addressed by the collection of a larger body of replication data that shows substantively the same pattern of performance.

Change in Operator Pool

As noted in the discussion of operator sub-populations, a considerable proportion of the MegaREG operator pool had no previous REG experience. Specifically, of the 24 operators who participated in the MegaREG experiments, only 11 had already participated in REG-type experiments. Thus, the difference between MegaREG and earlier REG performances might be due to the different population of contributing operators. Indeed, the results shown in Figure 5 and Table 3 show that the new and the previously experienced operators are indeed

very different populations. The difference between the high-density and low-density data is over five times as large for the operators with no previous REG experience. Interestingly, these new operators attain an independently significant \mathcal{D} in the direction of intention in the low-density data pooled across the entire experiment ($T = 2.1220$, $p = 0.017$, 1-tailed). (Measuring success in direction of intention is properly a one-tailed test.) It is, of course, impossible to determine from the current database whether this difference between the experienced and new operators arises simply because the new operators are a different set of people, whose idiosyncratic patterns of performance are by happenstance unusual relative to PEAR's earlier operator pool, or if it is the result of a systematic personal or psychological effect of having had previous REG experience.

Regardless of the difference between new and previously experienced operators, the operator-pool hypothesis is refuted by the fact that the experienced operators' high-density performance is independently significant ($T = -2.2975$, $p = 0.022$, 2-tailed) and significantly different from the REG benchmark effect ($T_d = -3.2027$, $p = 0.001$). Thus, operators who had participated in earlier REG experiments produced a MegaREG result consistently different from their earlier history. This may be due to some innate difference between the MegaREG experiment and other REG experiments, or it may be that an operator's performance in one experiment cannot be used reliably to predict performance in another. (Further evidence in support of this view is adduced in Jahn *et al.*, 2000, Table P.7, p. 538.)

Change in Task

MegaREG introduced a completely novel element to PEAR experiments by interspersing two distinct types of data in a fashion to which both operators and experimenters were blind. It was pointed out by several experimenters that this fundamentally changed the nature of the experiment, in that we were attempting to address two questions simultaneously: first, was there an anomalous effect; and second, did it differ between the two data types? Some teleological or observation-based models of anomalous phenomena would suggest that such a change in the basic analysis mode of an experiment might, in itself, induce a change in its outcome. As a subsidiary aspect of this issue, one also must note that in any experiment involving consciousness-related anomalies, it may not be possible to blind operators to an experimental condition by the mere absence of conventional sensory access to that condition.

Subject to the caveats of its smaller size and limited resolution, MegaMega shows that the double-blind interspersal is not the cause of the unexpected MegaREG outcome. With respect to intention, the result of MegaMega is indistinguishable from that of the high-density MegaREG data to which it corresponds, despite the absence of any double-blind interspersal in the MegaMega protocol. Whether or not MegaREG blinding was "successful" at

the level of the operators' unconscious where the potential for anomalies presumably resides (Jahn & Dunne, 2001), it does not seem to have had any effect on the experimental output, save perhaps in suppressing the non-intentional asymmetry that has appeared for female operators in other experiments.

A related question is the role of the *experimenter* blinding, which of course distinguishes the initial phase from the post-unblinding phase within MegaREG. The consistency between original and replication data indicates that unblinding the experimenters to the outcome did not have any major impact on the results.

Another consideration that falls under the category of "change in task" is the additional processing involved in both the low-density and the high-density trials. Of the two, the low-density trials are closer in concept and structure to the existing REG design: the discarding of most intervening bits is functionally equivalent to the temporally slower sampling used in the original REG experiment. In the high-density trials, on the other hand, a full two million bits were processed for each trial, but their collective outcome was then normalized for presentation in the same format as the low-density trials. Whether the difference between high-density and low-density outcomes arose from this difference in processing paths, rather than from the blinding and interspersal *per se*, is another "change in task" hypothesis that cannot be tested retrospectively. MegaMega does not resolve this issue, since it employed the same processing path as high-density MegaREG.

Processing Overload

Given the fact that the high-density data uniformly showed deviations contrary to the operator's intention, it is tempting to identify them as an actual dysfunction of whatever process is involved in creating intentional anomalies. The fact that this "dysfunction" is associated with an extremely high data rate adds to the temptation, since inadequate speed of processing sensory input has been implicated as a component in disorders ranging from schizophrenia to dyslexia. In other words, the high-density data might be contrary to intention simply because the operator's faculty for producing anomalous effects, whatever its nature, is overloaded by the tremendously high data rate, and rather than simply failing to function, operates erroneously.

In considering this speculation, we once again should bear in mind that the low-density MegaREG data are indistinguishable in their timing from the sampling pattern of the original REG. While the source continued to generate bits at the rate of 10 MHz, the low-density filtering operation accepted only every ten-thousandth bit. The remaining bits simply were discarded, neither presented to the operator nor recorded anywhere. From the operator's point of view, these discarded bits effectively did not exist. An operator presented with a low-density trial was seeing the sum of 200 bits gathered at a rate of $(10 \text{ MHz} / 10^4 = 1 \text{ kHz})$, exactly as in the earlier REG200 experiment. The contrast

between low- and high-density is thus exactly what one would expect if the intentional inversion is due to some form of processing overload.

A problem with this hypothesis is that the kilohertz sampling rate of the original REG and of the low-density MegaREG data already is considerably faster than most biological and neurological processes. Moreover, the REG2000 data show that reliable anomalous response is possible at bit rates of up to 10 kHz. It is not at all clear what physical or psychological capacities a human being might have that can operate at 10 kHz but break down dysfunctionally at 10 MHz. On the other hand, despite its *a priori* implausibility, the processing overload explanation at least has the virtue of being relatively easy to test. A series of experiments deployed at intermediate data rates could localize and verify a breakdown of intentional effort at a specific rate, if indeed this is the cause of the MegaREG high-density inversion.

Change in Environment

It is reasonable to expect that psychological factors would be relevant to the production of *consciousness*-related anomalies. Such factors include mode, attitude, prior beliefs, and environmental ambience (Braud *et al.*, 1995; Heath, 2000; Honorton & Barksdale, 1972; Polyani, 1983; White, 1976). Moreover, in this field, the possibility of experimenter effects must be considered seriously; the experimenter's own mood, attitude, and approach are part of the environment in which the operator generates data. While PEAR has striven to maintain a conducive and supportive atmosphere for the generation of anomalous results by operators, the history of the full range of REG experimentation suggests that control over this important parameter has been less than perfect (Jahn *et al.*, 2000). Psychological environment is, of course, a highly subjective parameter that is difficult even to specify systematically, much less to measure or control. Indeed, some aspects of the psychological environment may be inherently beyond experimental control. For example, an operator confronted with an REG experiment (or any experience) for the tenth or perhaps the hundredth time surely will not see it as fresh or novel, no matter what environment the experimenters seek to arrange. This should be taken into account when considering the stark contrast between new and previously experienced operators discussed above.

Nevertheless, some aspects of the MegaREG results are not plausibly attributable to environmental influences. The tremendous split between the high- and low-density performances appears between two classes of data generated by the same operator, at the same time, under the same environmental conditions. Since the conscious psychological situation does not differ between the two conditions, it can hardly be responsible for the difference in outcomes.

Conclusions

The results of three phases of MegaREG, and the companion MegaMega experiment, may be summarized as follows:

1. The experiments displayed a real and replicable anomalous effect in which the high-density data had an outcome contrary to intention.
2. This reversal with respect to intention refutes all straightforward scaling models.
3. The absolute value of the MegaREG effect is approximately three times (2.77 ± 0.69) larger on a per-trial basis, and so approximately 30 times smaller on a per-bit basis, than the REG200 effect.
4. A strong difference is found between performances of operators with and without previous REG experience. Both the primary intentional effect in each data density, and the difference between data densities, are much larger for the inexperienced operators.
5. The change in operator pool cannot be the source of the effect reversal, however, since both operator populations show it at a statistically significant level.
6. The difference between the MegaREG source and earlier REG sources is an implausible explanation for the difference between MegaREG and earlier REG experiments, because the high-density and low-density data generated on the same source also differ, and the low-density data are consistent with earlier REG experiments.
7. The double-blind interspersal of two data types is not the cause of the effect reversal.
8. Operator sensitivity to experimental task definition may be a possible explanation. This may be due to the increased processing load inherent in the higher data rate, or due to some aspect of the normalization process in high-density data.
9. The psychological environment, and/or experimenter expectations, may also be a factor in the overall production of a counter-intentional effect, but probably is not capable of explaining the split between data-density conditions.

Unfortunately, MegaREG is now a closed experiment. The noise source used has since suffered electronics failure, and the labor-intensive nature of repairing and maintaining the experiment renders impractical any effort to rehabilitate it at this time. Nevertheless, the replication phase has verified, at least, that the reversal phenomenon seen in the first phase was genuine, and that it is not sensitive to the experimenters' state of knowledge concerning the outcome. Further interpretation of this result remains elusive.

Acknowledgments

The Princeton Engineering Anomalies Research program gratefully acknowledges the financial support of the Institut für Grenzgebiete der Psychologie und Psychohygiene; The Lifebridge Foundation; Richard Adams; Laurance Rockefeller; and other private contributors. The experimental program described

herein was inspired and in large part designed by Michael Ibison, whose reports on the first phase have been referenced.

Appendix: Mathematical Demonstrations

Yield Unaffected by Normalization

A raw MegaREG trial value is (under the null hypothesis) a normally distributed random variable t with expectation $\langle t \rangle = 10^6$ and standard deviation $\sigma(t) = 1000/\sqrt{2} = 707.107$. Therefore the Z-score of such a trial is

$$Z(t) = \frac{t - \mu}{\sigma} = \frac{t - 1000000}{707.107}. \quad (\text{A.1})$$

The normalization process produces a normalized trial value $t' = 100 + (t - 1000000)/100$. This is referenced to the theoretical distribution for 200-bit trials, $\mu' = 100$, $\sigma' = \sqrt{50} = 7.07107$. The resulting Z-score is therefore

$$Z(t') = \frac{t' - 100}{7.07107} = \frac{(t - 1000000)/100}{7.07107} = \frac{t - 1000000}{707.107} = Z(t). \quad (\text{A.2})$$

Thus the Z-score of a normalized trial is the same as the Z-score of the corresponding raw trial. Statistical yield is unaffected by the normalization process.

Analysis Variable Derivations

Let the mean output level of the MegaREG source be μ , with standard deviation σ . Let h , l , and b be the means of three samples of MegaREG output; we make no assumption about the size of the samples. Let σ_h , σ_l , and σ_b be the standard deviation of the source during each of the respective sampling periods; while it would be convenient to assume that σ , like μ , is a constant of the device's operation we do not need this assumption. By hypothesis, $\langle h \rangle = \langle l \rangle = \langle b \rangle = \mu$. By definition, $\sigma_h^2 = \langle h^2 \rangle - \langle h \rangle^2 = \langle h^2 \rangle - \mu^2$; similar relations hold for l and b . We add the assumption that the three sample outcomes are mutually independent ($\langle hl \rangle = \langle h \rangle \langle l \rangle$ and likewise for the other two possible combinations).

We may now calculate directly the expected value and variance of the analysis variables \mathcal{D} and \mathcal{A} . For \mathcal{D} ,

$$\begin{aligned} \langle \mathcal{D} \rangle &= \left\langle \frac{h - l}{\sqrt{2}} \right\rangle = \frac{\mu - \mu}{\sqrt{2}} = 0; \\ \sigma^2(\mathcal{D}) &= \langle \mathcal{D}^2 \rangle - \langle \mathcal{D} \rangle^2 = \left\langle \frac{h^2 - 2hl + l^2}{2} \right\rangle - 0 = \frac{\langle h^2 \rangle + \langle l^2 \rangle - 2\langle hl \rangle}{2} \\ &= \frac{\langle h^2 \rangle + \langle l^2 \rangle - 2\langle h \rangle \langle l \rangle}{2} = \frac{\sigma_h^2 + \mu^2 + \sigma_l^2 + \mu^2 - 2\mu\mu}{2} \\ &= \frac{\sigma_h^2 + \sigma_l^2}{2}. \end{aligned} \quad (\text{A.3})$$

This establishes the relations of Equation 2 for \mathcal{D} . An exactly analogous derivation does the same for \mathcal{A} ; having shown the procedure involved we shall not take the space to show the full derivation.

We wish also to show that \mathcal{D} and \mathcal{A} are uncorrelated; that is, that they have zero covariance.

$$\begin{aligned} \text{Covar}[\mathcal{D}, \mathcal{A}] &= \langle \mathcal{D}\mathcal{A} \rangle - \langle \mathcal{D} \rangle \langle \mathcal{A} \rangle = \frac{\langle (h-l)(h+l-2b) \rangle}{\sqrt{12}} \\ &= \frac{(\langle h^2 \rangle - \langle l^2 \rangle) + (\langle hl \rangle - \langle hl \rangle) + (\langle 2bl \rangle - \langle 2bh \rangle)}{\sqrt{12}} = 0. \quad (\text{A.4}) \end{aligned}$$

Secular Drift: Empirical and Theoretical Solutions

The primary reason for adopting the measures \mathcal{D} and \mathcal{A} is the possibility of unknown departures from the theoretical distribution by the random source. Since one of the possible concerns is instability, or secular drift, in the distribution parameters, it seems obvious that the most reliable measures will be those derived from a single session. If the mean of a particular intention (h , l , or b) is the quantity being measured, the corresponding standard deviation (σ_h , σ_l , or σ_b) is best estimated as the standard error calculated from the trial-level standard deviation estimate: $\sigma = \text{s.e.} = s/\sqrt{n}$.

For concatenations of more than a single series, the pooled estimates of the mean and standard deviation in an intention could be contaminated by any parameter drift that might have taken place. If drift *within* a series is negligible, we may consider \mathcal{D} and \mathcal{A} to have expectation 0 when computed for a series (Eq. A.3), and to have known standard deviations in a given series, calculated by Eq. 2 from the individual standard errors of the intentional data. If secular drift is substantial within a series, however, the exact derivations showing $\langle \mathcal{D} \rangle = \langle \mathcal{A} \rangle = 0$ are invalid. Nonetheless, we may test for the possible existence of such short-term drift by examining the data within series. A change in the mean source output level during a series can be detected (and distinguished from possible intentional effects induced by the operator) by breaking the intentional data from a series into arbitrary subsets, and testing for variation between the means of these subsets in excess of variation expected on the basis of the trial-level variance. For this test, each intention's data were divided into ten segments. This corresponds to the run structure for series using short runs, and using the same number of segments in the series with long runs facilitates consistent evaluation of the outcome. The result of the test for the presence of secular drift is thus a χ^2 with 9 d.f. for each intention. All three intentions were examined, although, strictly speaking, only the baseline is a reasonable candidate for testing the presence of short-term drift, since the active operator-intention effects might show changes related to fatigue or other psychological responses. Also, the high-density and low-density data were tested separately for drift, producing six χ^2 values for each series.

TABLE A
Test for Secular Drift within Series

Intention	Low-density $\chi^2(p)$	High-density $\chi^2(p)$
High	1414 (0.081)	1399 (0.132)
Baseline	1288 (0.847)	1409 (0.096)
Low	1385 (0.197)	1386 (0.192)

The composite result of this test for the existence of within-series drift can be computed by summing the χ^2 values across the 149 series of original and replication data: the sum of 149 χ^2 s, each with 9 d.f., is a single χ^2 with 1341 d.f. Across the three intentions and two densities, these χ^2 values (and associated p -values) are given in Table A. Even in the active intentions, none of these figures attain statistical significance. We may conclude that any short-term secular drift of the noise source taking place within individual series is too small for statistical detection in this database and hence may be disregarded in the subsequent analyses.

The data can be protected from longer-term secular drift effects between series by calculating \mathcal{D} and \mathcal{A} at the series level, and combining these parameters, rather than the raw data, when multi-series concatenations are needed. The combination formulae for n values m_i , $i = 1, \dots, n$, with corresponding standard error estimates s_i , are:

$$m = \frac{\sum_{i=1}^n m_i/s_i^2}{\sum_{i=1}^n 1/s_i^2}; \quad \sigma^2[m] = \frac{1}{\sum_{i=1}^n 1/s_i^2}. \quad (\text{A.5})$$

References

- Braud, W., Shafer, D., McNeill, K., & Guerra, V. (1995). Attention focusing facilitated through remote mental interaction. *Journal of the American Society for Psychological Research*, *89*, 103–115.
- Dobyens, Y. H. (2000). Overview of several theoretical models on PEAR data. *Journal of Scientific Exploration*, *14*, 163–194.
- Dobyens, Y. H., Dunne, B. J., Jahn, R. G., & Nelson, R. D. (2002). *The MegaREG Experiment: Replication and Interpretation (revised edition)*. Technical Note PEAR 2002.03.
- Dunne, B. J. (1991). *Co-Operator Experiments with an REG Device*. Technical Note PEAR 91005.
- Dunne, B. J. (1998). Gender differences in human/machine anomalies. *Journal of Scientific Exploration*, *12*, 3–55.
- Heath, P. R. (2000). The PK zone: A phenomenological study. *Journal of Parapsychology*, *64*, 53–72.
- Honorton, C., & Barksdale, W. (1972). PK performance with waking suggestions for muscle tension versus relaxation. *Journal of the American Society for Psychological Research*, *66*, 208–214.
- Ibison, M. (1997). *Evidence That Anomalous Statistical Influence Depends on the Details of the Random Process*. Technical Note PEAR 97007.
- Ibison, M. (1998). Evidence that anomalous statistical influence depends on the details of the random process. *Journal of Scientific Exploration*, *12*, 407–423.
- Jahn, R. G., Dobyens, Y. H., & Dunne, B. J. (1991). Count population profiles in engineering anomalies experiments. *Journal of Scientific Exploration*, *5*, 205–232.
- Jahn, R. G., & Dunne, B. J. (2001). A modular model of mind/matter manifestation (M^5). *Journal of Scientific Exploration*, *15*, 299–329.
- Jahn, R., Dunne, B., Bradish, G., Dobyens, Y., Lettieri, A., Nelson, R., Mischo, J., Boller, E., Bösch,

- H., Vaitl, D., Houtkooper, J., & Walter, B. (2000). Mind/Machine Interaction Consortium: PortREG replication experiments. *Journal of Scientific Exploration, 14*, 499–555.
- Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H., & Bradish, G. J. (1997). Correlations of random binary sequences with pre-stated operator intention: A review of a 12-year program. *Journal of Scientific Exploration, 11*, 345–367.
- Nelson, R. D., Jahn, R. G., Dobyns, Y. H., & Dunne, B. J. (2000). Contributions to variance in REG experiments: ANOVA models and specialized subsidiary analyses. *Journal of Scientific Exploration, 14*, 73–89.
- Polyani, M. (1983). *The Tacit Dimension*. Gloucester, MA: P. Smith.
- Radin, D. I., & Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics, 19*, 1499–1514.
- Schmidt, H. (1970a). Quantum mechanical random number generator. *Journal of Applied Physics, 41*, 462–468.
- Schmidt, H. (1970b). A PK test with electronic equipment. *Journal of Parapsychology, 34*, 175–181.
- Schmidt, H. (1974). Comparison of PK action on two different random number generators. *Journal of Parapsychology, 38*, 47–55.
- White, R. A. (1976). The influences of persons other than the experimenter on the subjects' scores on psi experiments. *Journal of the American Society for Psychical Research, 70*, 332–370.