Journal of Scientific Exploration, Vol. 14, No. 1, pp. 73–89, 2000 0892-3310/00

Contributions to Variance in REG Experiments: ANOVA Models and Specialized Subsidiary Analyses

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

Princeton Engineering Anomalies Research
School of Engineering/Applied Science
Princeton University, Princeton NJ 08544

Abstract—Judicious application of a complementary set of sophisticated analytic techniques to large databases from human/machine anomalous interaction experiments can extract subtle structural features that might elude more simplistic analyses. The combination of a multi-factor analysis of variance (ANOVA) with various subsidiary, *ad hoc* approaches suggested by the ANOVA or directly by the data, can establish an instructive hierarchy of salient physical and subjective parameters and illuminate some of their specific details. In this particular study, the dominant finding is a significant correlation of anomalous effects with prescribed intentions of the human operators, compounded of small contributions from many individuals across many experimental conditions. The grand concatenation, which includes all combinations of successful and unsuccessful parameters or conditions, shows a chance probability for this correlation with intention on the order of 10^{-4}. The effect apparently is confined to non-deterministic devices; *i.e.*, deterministic pseudo-random sources show no overall effect. The correlation with intention for non-deterministic sources alone has a chance probability of 10^{-6}. Beyond operator intention, most of the other technical, procedural, and subjective parameters explored show unimpressive contributions to the overall variance, with a few notable exceptions that are clarified in the subsidiary analyses. For example, individual differences among operators are indicated, but there is a relatively normal distribution of effect sizes, within which a few participants are distinguished by consistent achievement over large databases. The temporal development of effect sizes shows a consistent pattern of initial success that declines but then recovers. There is essentially no evidence for a dependence of effect size on spatial or temporal separation, supporting other indications that ordinary physical variables have little impact on the anomalous interactions. In sum, although the composite ANOVA models explain less than 1% of total variance, implying very small and subtle effects, the analysis provides strong evidence that the anomalies are statistically robust; they are not due to chance fluctuations, but are demonstrably correlated with definable subjective factors.

Keywords: anomalies — ANOVA — consciousness — electronic random event generator — mind/machine interactions — models — REG — RNG

Introduction

Beginning about four decades ago, electronic random event generators or random number generators (REG or RNG) have been used in a wide range of laboratory experiments designed to test the hypothesis that human consciousness might interact directly with labile physical systems. The results provide clear statistical evidence that the behavior of these devices deviates from chance expectation in correlation with the pre-defined intentions of participants in the experiments (Radin and Nelson, 1989). In 1979, the Princeton Engineering Anomalies Research Laboratory (PEAR) began collecting large databases in REG experiments using particularly rigorous controls and a variety of optional parameters to assess the character and replicability of such anomalous mind/machine interactions. Over a 12-year period of primary investigation, ten physical and psychological conditions were examined as possible mediating variables in the experimental results. A number of extensions and variations of the basic protocol were explored, using different REG sources as well as a selection of other physical systems, the performance of which was dependent on some form of random process. The experiments had accompanying calibrations which confirmed that the random sources were of high quality, producing data that consistently conformed to theoretical expectations in non-experimental conditions. In the active experiments, however, the data sequences and distributions were significantly correlated with the experimental variables, especially the operators' intentions to shift the means of the REG output distributions, and showed structure that could not be accounted for by chance fluctuations.

This paper presents a compact summary of a formal analysis of variance (ANOVA), and a number of subsidiary *ad hoc* analyses of the results from these REG studies, comprising 1338 replications of the experiment, including a small set of variations on the basic protocol.

Equipment

The PEAR program has used three generations of random event generators, utilizing different primary sources of white noise but maintaining important common features of design. An original benchmark experiment employed a commercial random source sold by Elgenco, Inc., the core of which is proprietary. Elgenco's engineering staff describe this module as solid state junctions with precision preamplifiers, implying processes that rely on quantum tunneling to produce an unpredictable, broad-spectrum white noise in the form of low-amplitude voltage fluctuations (Nelson, Bradish, and Dobyns, 1989). A much simpler and more compact portable REG was based on Johnson noise in resistors, or so-called thermal noise, which also is a quantum-level phenomenon that produces a well-behaved, broad-spectrum fluctuation (Nelson, Bradish, and Dobyns, 1992). A later-generation device called the PEAR Micro-REG used a field effect transistor (FET) for the primary noise source, again relying on quantum tunneling to provide uncorrelated fundamental events that compound to an unpredictable voltage fluctuation.

In all cases, the design begins with some white-noise frequency distribution. For example, the benchmark REG presents a flat spectrum $\pm$ 1 db from

50 Hz to 20 kHz. A low-end cutoff at 1000 Hz attenuates frequencies at and below the data-sampling rate. This filtering, followed by appropriate amplification and clipping, produces an approximately rectangular wave train with unpredictable temporal spacing. Gated sampling, typically at a constant 1-kHz rate, yields a regularly spaced sequence of random bits, suitable for rapid counting. Other sources have been constructed that allow higher sampling rates, up to 2 MHz (Ibison, 1998), but this paper summarizes data from the standard unit only. Analog and digital processes are isolated by temporally alternating these operations to exclude contamination of the analog noise train by the digital pulses. To eliminate biases of the mean that might arise from such environmental stresses as temperature change or component aging, an exclusive or (XOR) mask is applied to the digital data stream. This is either a regularly alternating 1/0 pattern or a more complex mask comprising a randomly ordered array of all 8-bit bytes with equal occurrence of 1/0. The latter procedure also excludes all short-lag bit-to-bit and byte-to-byte auto-correlations. Finally, data for the experiments are presented and recorded in "trials" that are the sum of *N* samples (typically 200 bits) from the primary sequence, thus further mitigating any residual short-lag auto-correlations. The final output of the benchmark REG thus is a sequence of conditioned bits, and in the later devices, of bytes presented to the computer's serial port, which then are formed into a sequence of trials, usually generated at approximately 1 per second. Calibrations on all of the devices closely conform to statistical expectations for the mean, variance, skew, and kurtosis of the accumulated count distributions, and expectations for time-series of independent events (Nelson, Bradish and Dobyns, 1989; Nelson, 1993; Nelson *et al.*, 1997).

Experimental Design

The basic experimental designs also embody protocol-level protection against artifactual sources of apparent effect. Following a "tripolar" protocol, participants generate data under three conditions of pre-specified intention, namely to achieve high (HI) or low (LO) mean values, or to generate baseline (BL) data. With the exception of the intention held in the participant's mind, which is pre-recorded in computer files, these three conditions are otherwise the same; all potentially influential variables are maintained constant within an experimental session or series.

In addition to the primary variable of tripolar intention, a number of secondary parameters are available as options that can be explored in separate sessions, and assessed as factors that may contribute to the experimental outcomes. These include:

1. Human variables such as the identity of the individual operators (participants), their gender, and whether they are "prolific," *i.e.*, have performed sufficient replications of the experiment to permit robust comparisons. In this category we also include the replication number or

serial position of the session as a factor that reflects operator experience.

2. Physical variables such as the different noise sources, including not only the true random sources described earlier, but also various hardware and algorithmic pseudo-random generators, designated as non-deterministic and deterministic sources, respectively.
3. Operational variables, including spatial separation of the operator from the machine (up to thousands of miles) and separations in time (up to several hours or in a few cases, one to four days); information density (bits per second); the number of trials in automatically sequenced "runs;" the instruction mode (volitional or instructed); and the type of feedback provided to the operator.

Analysis of Variance

The benchmark REG database was accumulated over a period of 12 years, with contributions from 108 individual operators, 30 of whom met our criterion for prolific operators by generating a minimum of 10,000 trials per condition (the equivalent of 10 experimental series or replications). The primary database of 1262 independent replications, comprising a total of 5.6 million 200-bit trials, was analyzed by a regression-based ANOVA, previously reported (Nelson *et al.*, 1991). Since this phase was completed, two smaller data sets have been added, for a total of 1338 replications, and new versions of two factors (the random source variable is simplified; operator identification is now enhanced with gender specification) included in the original analysis have been defined to help understand the results of the original analysis and those of other special-purpose, detailed analyses. Although they are defined *a posteriori,* they provide legitimate assessments of the questions they represent, within the context of the full complex of potential influences in this experiment. The discussions that follow thus rely on the original analysis, supplemented by additional information derived from the expanded assessment using the new factor definitions.

The REG database is complex, involving nine analytical factors with two to five levels, and a tenth (operators) with more than 100. The analytical matrix has unbalanced cells, requiring that the ANOVA be based on multiple regression modeling, employing a "model-comparison" procedure for partitioning the regression sum of squares. In addition to the comprehensive model and a similar analysis addressing only the contributions of prolific operators, a number of smaller models have been used to examine the details of variance contributions in particular subsets and in individual operator databases. Details of these may be found in the earlier technical report (Nelson *et al.*, 1991).

The formal models indicate that anomalous effects appear as small statistical signals in a background of noise; in most cases the amount of variance explained is on the order of one percent or less. Hence, it is only through the accumulation of large databases that these small effects can be examined.

Full Database Analysis

The tables, figures, and comments in this section summarize the major results of the original 1991 analysis of variance (Nelson *et al.*, 1991) based on models for all (1262) replications, the prolific operator subset, and a selection of restricted subsets relevant to particular questions. Each regression model is evaluated in terms of its sum of squares (*SSR*), degrees of freedom (*df*), associated *p*-value (*p*), and proportion of total variance explained by the parameters used in the model (R^2). Following this, a partitioning of the regression sum of squares reveals the contributions due to intention and to each of the secondary parameters. The latter combine the main effect and the interaction with intention for each parameter, and include an "unaccounted" entry that indicates the average over- or underestimation of factor contributions resulting from the model-comparison approach to the breakdown of the regression sum of squares. This is on the order of 1% or less of the total residual variance and is not significant. Tables 1–6 all have the same format, showing the composite sum of squares, degrees of freedom, *F*-ratio, and a *p*-value indicating the significance of each factor's contribution, where "*p*-value" refers to the chance of an outcome more extreme than the observed one, assuming the null hypothesis. In subsequent discussions, the terms "suggestive" or "marginal" refer to *p*-values between 0.10 and 0.05, and "significant" to those less than 0.05. Figures 1–4 illustrate the overall effect of parameters, usually by displaying the mean shift, or effect size, as a function of intention, with one-sigma error bars. The unit for the effect size is the equivalent number of bits per 50 trials, and corresponds to the number of excess bits per 10,000 binary events.

The model for the entire data set, with the decomposition of the total sum of squares due to the experimental factors, is presented in Table 1. The corresponding overall mean shifts are shown in Figure 1. All data and parameters

TABLE 1
All Data (All Parameters Except Operators)

Model: $SSR = 199220$; $df = 50, 113703$; $p = 0.0050$; $R^2 = 0.00070$

Parameter	Sum of Squares	*df*	*F*-Ratio	Probability
Intention	41875	2	8.349	2.4×10^{-4}
Device	32537	6	2.163	0.043
Location	28618	6	1.902	0.076
Protocol	9954	6	0.662	0.68
Series	40934	12	1.360	0.18
Runlength	11748	6	0.780	0.58
Assignment	4318	3	0.574	0.63
Control	7730	3	1.028	0.38
Feedback	19475	6	1.294	0.26
Unaccounted	2032			

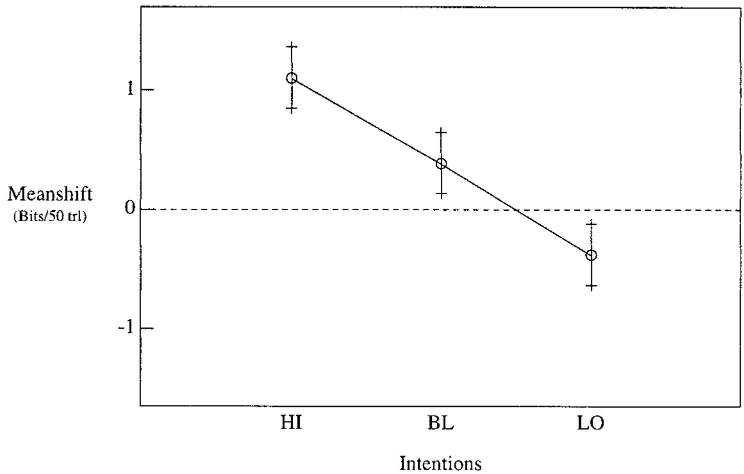

Fig. 1. Correlation of mean shift with intention. All data are from the comprehensive model including all factors except operators. Units for the mean shift are the number of bits per 50 trials, which is equivalent to the number of bits per 10,000. Point estimates (circles) with one-sigma error bars.

are included, with the exception of the operators factor, whose interactions with intention comprise more than 300 degrees of freedom, entailing calculations that exceed the computational system limits. This factor is addressed later in the prolific operators model.

The regression model for the full database across all operators and parameters is significant, clearly indicating that the combined influence of the experimental parameters adds information to the nominally random distribution. The model explains less than 0.1% of the variance, however, which is consistent with the very small ratio of anomalous effect to stochastic "noise." We will see in subsequent analyses that certain well-defined subsets have an R^2 an order of magnitude larger, but none of the models explain much more than 1% of variance.

As predicted, intention is the primary contributor to the regression, with a mean squared deviation (and F-ratio) four times that of the next largest parameter, and, as the figure shows, the relationship between intention and meanshift is consistent with the experimental hypothesis. While HI and LO are properly displaced relative to BL, the meanshift relative to theoretical expectation is greater for HI than LO. This imbalance appears consistently in various subsets but the differences are not statistically significant. The BL data also show a tendency for high deviations that are fairly consistent, although not significant, in curious contrast to the calibrations, which properly do not exhibit any consistent trends (Nelson, 1993). "Device" (type of random source) and "location" are significant and marginal secondary parameters, respectively, suggesting a need for separate examination of the various levels of each. Briefly, such subsidiary analyses show that results with the diode source and with a non-deterministic shift-register-based device are similar to each other, but differ from those with the deterministic or algorithmic pseudo-random devices. The shift-register device originally was classified as pseudo-random, but subsequently was determined to be a combination of pseudo-random and truly random components, hence qualifying as a non-deterministic source (Jahn *et al.*, 1997). With respect to "location" the subsidiary analysis shows a second-order interaction: local and remote data are similar, while the "B" location (operator in the next room, delayed feedback) differs, but the latter effect is driven by a confounding interaction with the device type. This will be discussed further, but we should note in this context that the database for the "B" location is comparatively small and hence any inferences must be tentative. To examine the effect of large spatial separations on the anomalous effects more thoroughly, a special-purpose analysis using linear regression was applied. The results show no trend in scoring associated with increasing separation of operator from machine, up to several thousands of miles (Dunne and Jahn, 1992).

In the full model, "assignment" (whether the instruction was random or volitionally chosen) and "control" (automatic *vs.* manual sequencing of trials) are not represented precisely because 16 early series mixed these parameters. When these series are excluded, "assignment" is associated with a p-value of 0.52, and "control" with a p-value of 0.29, hence neither factor is a contributor in the grand concatenation. The "protocol" (the experiment's length and purpose, particular questions, *etc.*), "series" (series position), "runlength," and "feedback" factors are all non-significant in the main analysis. For "series," however, other detailed assessments indicate a clear, albeit complex, non-linear structure (Dunne *et al.*, 1994) which will be discussed later.

In qualification tests of the multiple regression modeling procedure, the corresponding analysis applied to arbitrarily assigned calibration data in the full database model yielded a non-significant regression, and parameter contributions that were well within chance variation: one factor had a marginal p-value (0.063), but none reached the nominal significance level (Nelson *et al.*, 1991). Thus, the appearance of structure in the experimental database cannot be attributed to artifacts of the modeling procedure.

Prolific Operator Results

Table 2 and Figure 2 show the results for a subset of the full original database consisting of 1060 series produced by the 30 "prolific" operators, each of whom generated 10,000 or more trials per intention. In this model, which constitutes approximately 80% of the full database, the operator factor can be assessed, and because the various parameter levels are represented more evenly, the results may be interpreted with higher confidence. This prolific operator model is highly significant and explains more than twice as much of the variance (0.2%) as the full model, suggesting that the inclusion of both multiple

TABLE 2
Prolific Operators (All Data and All Parameters)

Model: $SSR = 499986$; $df = 137, 97135$; $p = 4.4\times10^{-4}$; $R^2 = .0020$

Parameter	Sum of Squares	*df*	*F*-Ratio	Probability
Intention	44215	2	8.790	1.5×10^{-4}
Operators	298546	87	1.364	0.014
Device	42382	6	2.809	0.010
Location	36848	6	2.442	0.023
Protocol	14755	6	0.978	0.44
Series	24260	12	0.804	0.65
Runlength	10263	6	0.680	0.67
Assignment	7194	3	0.954	0.41
Control	4122	3	0.546	0.65
Feedback	27343	6	1.812	0.092
Unaccounted	−9942			

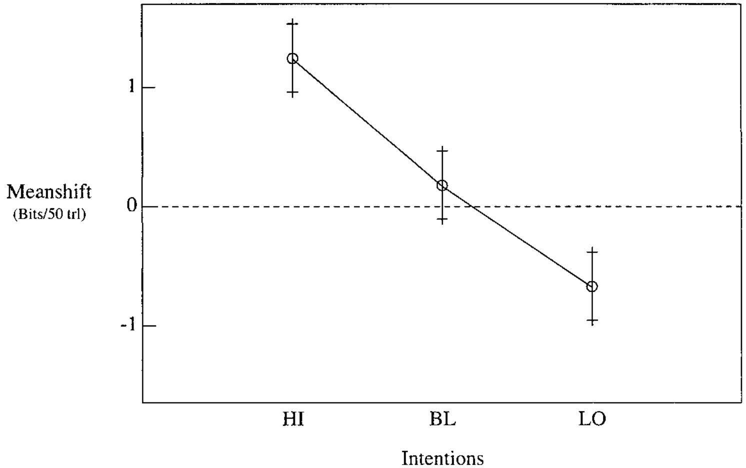

Fig. 2. Correlation of mean shift with intention. Data are from the prolific operators model including all factors. Mean shift units are the number of bits per 50 trials, which is equivalent to the number of bits per 10,000. Point estimates (circles) with one-sigma error bars.

replications by individuals and a factor that represents individual differences may serve to clarify the effects of the parameters.

Again, intention is highly significant and is the primary contributor to the regression. The operator parameter also is significant, indicating individual differences in performance and the need to consider each operator's data separately to assess individual responses to parameters. Such analyses have been detailed in the earlier technical report (Nelson *et al.*, 1991), and will be considered further in discussion of the updated ANOVA. Figure 3 shows the results of a cluster analysis of the intention-linked performance of the 30 prolific operators. There are three well-defined clusters: one for success in the direction of intention (HI – LO), one for a small number of operators with

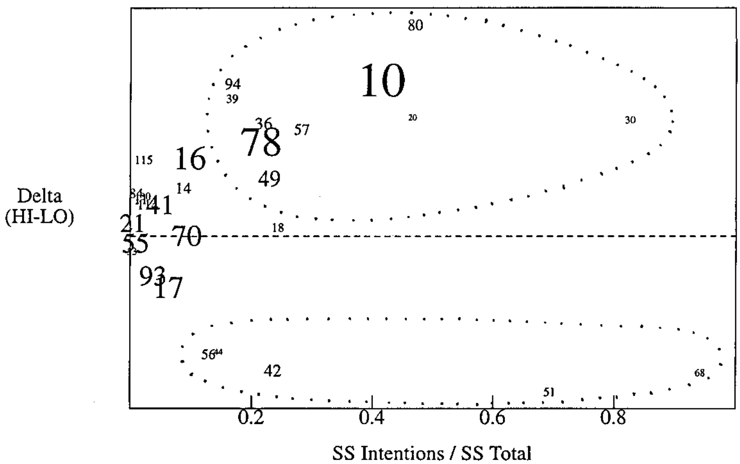

Fig. 3. Clustering of prolific operator effect sizes. The difference between HI and LO mean shift (arbitrary units) is plotted against the proportion of the sum of squares attributable to intention. Operator numbers are printed in a point size proportional to the operator's database size. The computed clusters are indicated by dotted lines around a group with large effect sizes in the intended direction (upper group) and opposite to intention (lower group). The remaining operators did not generate effects correlated with intention.

large effects in the direction opposite to intention, and a third cluster for whom intention was not a significant contributor to the regression model.

Both device and location appear as important secondary parameters, similar to the grand concatenation shown in Table 1, but for location we must again apply the previously mentioned caveats concerning the device interaction stemming from the small "B" data set. Feedback shows a marginal contribution, and examination of the subset means suggests, surprisingly, that the simple digital feedback and non-feedback conditions produce somewhat higher scores than the more engaging and informative graphic feedback mode while detailed examination shows this to be driven largely by early trials, which had larger effect sizes but which were done only with digital feedback. Subsequent studies suggest that feedback may indeed be an important parameter.

Although the series position contribution to the multiple regression is non-significant, the assessment of subset means reveals a complex pattern of results as a function of series or replication number. As shown in Figure 4, the highly significant composite result in the first series declines to non-significance, and then recovers to a significant effect in later series. The figure shows that this non-linear progression of effect size is superimposed on a weakly defined linear trend, thus explaining why the contribution of this parameter to the simple regression model is modest.

This pattern has been the subject of a more specific and detailed regression analysis (Dunne *et al.*, 1994) that confirms and extends the indication of a strong influence of series position (corresponding to developing operator

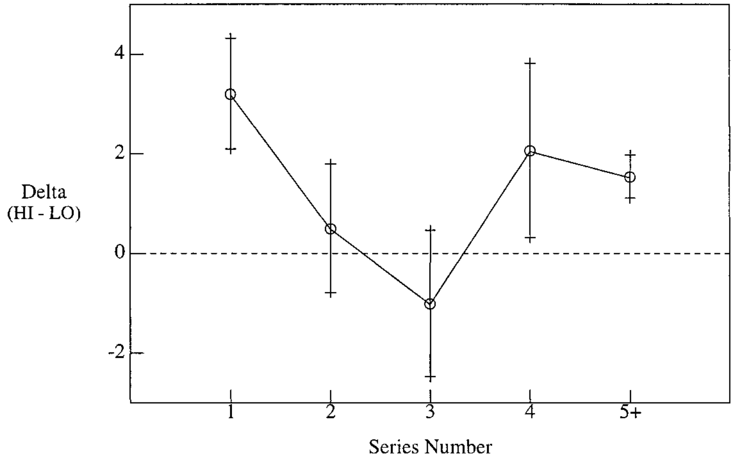

Fig. 4. Sequential development of effect size as a function of series position. Units of effect size are bits per 50 trials, and the differences between the HI and LO intentions for each series are plotted as point estimates (circles), with one-sigma error bars. The point labeled 5+ includes series 5 and all subsequent series.

experience), despite the modest contribution from this factor in the primary ANOVA model. The subsidiary analysis reveals that a similar pattern of strong early performance followed by a decline and subsequent recovery obtains in both the high and low intentions. The quadratic component of the regression is significant (p = 0.016) for the high–low difference, while the linear component is negligible. No such pattern is evident in the baseline data.

Most of the other secondary parameters in the ANOVA model are not overall contributors across operators, but a number of individuals respond differently to the experimental variations. Detailed analyses of the associated patterns are beyond the scope of this paper, but may be found in the technical report (Nelson *et al.*, 1991). The unaccounted variance in the prolific operator model is negative, indicating average overestimation of contributions, but this is inconsequential in magnitude.

Again a corresponding analysis using arbitrarily assigned calibration data yields a non-significant regression and no evidence of non-chance variation. One parameter has a marginal p-value (0.086), but it is a factor different from that with a marginal p-value in the full database model, giving further evidence that calibration data show only normal chance fluctuations, and that the modeling procedure does not produce artifactual indications of parameter effects.

ANOVA Update

The specialized analyses of the REG database using the *ad hoc* tools and perspectives mentioned above revealed some influential variables not included in the original, 1991 analysis, and prompted the development of corresponding parameters to be included in an update of the analysis of variance. Of course this is an *a posteriori* procedure, in the sense that the questions arose out of specific analyses of the data, but the new factors are reasonable extensions or modifications of those in the original ANOVA.

In this updated analysis, a factor called "gender" is represented with three levels: male (59 operators), female (50 operators), and co-operators (18 pairs). Although Dunne found instructive, significant differences in the average effects of men and women (Dunne, 1991), this factor is not a significant contributor in either the new overall model (see Table 3) or the corresponding prolific operators model. The finding is consistent with Dunne's observation that the compounded effects (composite Z-scores) do not differ significantly for men and women, although their average effect sizes do differ (Dunne, 1998). This is explained by the disproportionate contributions in very large databases by three female operators whose results differ significantly from the other forty-seven.

The co-operator data were not included in the 1991 analysis, but they have been assessed directly in a specialized, *ad hoc* analysis (Dunne, 1991). In particular, the effect sizes have been compared with those of single operators and appear to be larger by a factor of two or more, depending on the composition of the pairs. Despite the comparatively large mean shift for the co-operators, this database is too small relative to those for the individual male and female operators to produce a significant difference across the gender factor. We made no attempt to confirm the striking "bonded-pair" results found by Dunne because this would require a further dilution of cell populations in the co-operator subset of the database.

The redefinition of the device parameter as a deterministic or non-deterministic "source," and the inclusion of "gender," have two other notable effects. The previous indication that location was a significant or marginal

TABLE 3
Revised Model, New Parameter Definitions:
Source (Random *vs.* Deterministic), Gender (Female, Male, Co-Operators)

Model: $SSR = 204391$; $df = 56, 120987$; $p = 0.0147$; $R^2 = 0.00067$

Parameter	Sum of Squares	df	F-Ratio	Probability
Intention	43999	2	8.772	1.5×10^{-4}
Source	24655	3	3.277	0.020
Location	20148	9	0.893	0.53
Protocol	16297	6	1.083	0.37
Series	50770	12	1.687	0.063
Runlength	13152	6	0.874	0.51
Assignment	4787	3	0.636	0.59
Control	7967	3	1.059	0.37
Feedback	11293	6	0.751	0.61
Gender	3296	6	0.219	0.97
Unaccounted	8027			

parameter is mitigated here, as a result of clarifying the device distinction. Secondly, due to clearer definition of the device influences, as well as possible effects of gender differences, the series position parameter now is marginally significant.

A separate, special-purpose definition of the gender factor using four levels includes, in addition, a special category that segregates the contributions of three high-performing female operators with moderate to large effect sizes and extremely large databases, resulting in disproportionate contributions to the overall effect. Their operator numbers (10, 78, 80) are notable in Figure 3, where the font size is proportional to the database size. This four-level "gender" factor is obviously defined *a posteriori*, and should be regarded as a means to confirm findings of other, independent analyses, (Jahn *et al.*, 1997), within the context of the full analysis of variance. The full ANOVA model using this four-level parameter shows it to be a highly significant contributor ($p = 3.4 \times 10^{-4}$), indicating that results for the small, selected group of operators clearly differ from the general pattern (see Table 4 and Figure 3).

Although this conclusion must be tempered by the *a posteriori* nature of the analysis, it is clear that the combination of very large databases with moderate, positive effects contributes powerfully to the experimental outcome. Understanding the contributions of this source of variance may help to interpret other aspects of the experiment; it underscores the importance of large individual databases, where interactions with other variables are not confounded by significant individual differences.

The updated regression models also include a newly defined device factor, called "source," with two levels: non-deterministic and deterministic. The former comprises the Elgenco-based REG and the first version of a hardware "pseudo-random" source, which was, in fact, non-deterministic because it employed randomly varying shift-register steps. A revised hardware pseudo-random source subsequently was developed, as well as an algorithmic pseudo-

TABLE 4
Revised Model, All Data: Contribution of High Performers

Model: $SSR = 278012$; $df = 59, 120984$; $p = 0.000051$; $R^2 = 0.00092$

Parameter	Sum of Squares	*df*	*F*-Ratio	Probability
Intention	43999	2	8.774	1.5×10^{-4}
Source	25386	3	3.375	0.017
Location	23026	9	1.020	0.42
Protocol	14683	6	0.976	0.44
Series	47974	12	1.594	0.085
Runlength	11007	6	0.732	0.62
Assignment	7160	3	0.952	0.41
Control	11313	3	1.504	0.21
Feedback	14900	6	0.990	0.43
Gender+hi−perf	76917	9	3.409	3.4×10^{-4}
Unaccounted	8026			

random source. The latter two are properly deterministic, and data generated using them are compared with those from the truly random, non-deterministic devices by means of the new source factor. The updated models show this parameter to be significant, with a p-value of 0.020 in the grand, overall regression, and $p = 0.0071$ in the prolific operator model.

When the non-deterministic sources, also shown in other analyses to yield no anomalous effect (Jahn *et al.*, 1997), are excluded (see Table 5), the explanatory power of the model increases greatly to nearly 1% of the variance, and the significance of the intention factor increases to 10^{-6}. No other factor except that representing series position is a prominent contributor to this model; without the relatively noisy deterministic source data, the series parameter becomes significant. It is also worth noting that in this model location shows no suggestion of differentiation, confirming other evidence that anomalous effects are not a function of spatial separation.

Table 6 presents a model restricted to the deterministic sources alone, where we find no indication that the intention factor contributes to an

TABLE 5
Revised Model, Non-Deterministic Sources Only

Model: $SSR = 197238$; $df = 53, 90210$; $p = 0.012$; $R^2 = 0.00087$

Parameter	Sum of Squares	df	F-Ratio	Probability
Intention	69428	2	13.861	1.0×10^{-6}
Location	6492	9	0.288	0.98
Protocol	23379	6	1.556	0.16
Series	55234	12	1.838	0.037
Runlength	23402	6	1.557	0.16
Assignment	9045	3	1.204	0.31
Control	5509	3	0.733	0.53
Feedback	8253	6	0.549	0.77
Gender	3454	6	0.230	0.97
Unaccounted	957			

TABLE 6
Revised Model, Deterministic Sources Only

Model: $SSR = 128858$; $df = 47, 30732$; $p = 0.313$; $R^2 = 0.0017$

Parameter	Sum of Squares	df	F-Ratio	Probability
Intention	1914	2	0.380	0.68
Location	64171	9	2.832	0.0025
Series	14753	12	0.488	0.923
Runlength	8999	6	0.596	0.73
Assignment	3723	3	0.493	0.687
Control	18599	3	2.463	0.061
Feedback	15140	6	1.002	0.42
Gender	14027	6	0.929	0.47
Unaccounted	2467			

explanation of variance. Only the location factor is significant in this case, while the regression model itself is non-significant. Thus, the indications of differentiation by location can be attributed to the large, though inconsistent effects in the small "B" and "C" subsets of the deterministic data.

Combining these two specialized questions in a single analysis, we generate a model that includes the exploratory parameter segregating the three high-performing operators, and excludes data from deterministic sources (which show no effect). The variance explained in this model increases to well over 1% (see Table 7). Intention is the primary contributor, followed by the four-level gender factor, the explanatory power of which is about one-third that of intention. Of the remaining parameters, only series position appears as a marginally significant contributor.

Conclusions

The most important finding in both the original and the updated analyses is a significant correlation of outcome with the pre-assigned intentions, compounded largely of small contributions from many individual operators across most of the experimental conditions, but with a disproportionate contribution from the three high-performing operators. Depending upon the particular subset of the large and complex database, the statistical significance ranges up to a few parts per million, with the grand concatenation, which includes all combinations of successful and unsuccessful parameters or conditions, showing a probability for the correlation with intention on the order of 2×10^{-4}.

These comprehensive ANOVA models provide compact summaries of the major findings of the REG experiments, with the combined effect of all measured parameters taken into account. In conjunction with the complementary analyses addressing specific questions, this approach leads to a number of well-defined conclusions:

TABLE 7
Revised Model:
Segregation of Non-Deterministic Sources and Contribution of High Performers

Model: $SSR = 291671$; $df = 56, 90207$; $p = 4.0 \times 10^{-6}$; $R^2 = 0.0013$

Parameter	Sum of Squares	*df*	*F*-Ratio	Probability
Intention	69428	2	13.866	1.0×10^{-6}
Location	7494	9	0.333	0.96
Protocol	23698	6	1.578	0.15
Series	47481	12	1.581	0.089
Runlength	21943	6	1.461	0.19
Assignment	11376	3	1.515	0.21
Control	9455	3	1.259	0.29
Feedback	9965	6	0.663	0.68
Gend+hi−perf	97887	9	4.344	1.1×10^{-5}
Unaccounted	−7057			

1. Overall, the correlation with intention is a small-magnitude effect, equivalent to a distribution mean shift of about 1 part in 10,000, ranging up to an order of magnitude larger in certain subsets. This finding differs little between the original and updated analyses ($p = 0.00015$, $p = 0.00024$, respectively). This result also is similar to that in Jahn *et al.* (1997) for the benchmark REG ($p = 0.00007$), although a direct comparison is not appropriate since the ANOVA result considers the structuring effect of all three intentions, while the standard analysis regards only the differential between the HI and LO conditions.
2. The broad generality of the finding across most of the different combinations of parameters suggests a mechanism that operates at a very fundamental level. For example, a separate, specialized analysis indicates that the anomalous effect can be modeled most simply as an alteration of the fundamental binary probability of the random events (Jahn, Dobyns, and Dunne, 1991).
3. The effect is apparently confined to non-deterministic devices. The updated version of the ANOVA confirms that deterministic pseudo-random sources do not change behavior in correlation with the operators' intentions, while non-deterministic random sources incontrovertibly do so. The full regression model indicates a significant contribution of the "source" parameter, and a model limited to data taken with non-deterministic sources is highly significant (4.0×10^{-6}) with the intention parameter also significant (1.0×10^{-6}), while the corresponding model for data from deterministic sources alone is not significant, nor is the intention parameter. (We note, however, that location is significant and control is marginal in the latter model [$p = 0.002$ and 0.061, respectively], and since other researchers have reported effects with pseudo-random sources (Radin and Nelson, 1989), a strong conclusion discounting deterministic sources would be premature.)
4. Although there are individual differences, there is a relatively normal distribution of effect sizes across individuals, with no indications of outliers indicating special performance in the sense that only certain "gifted" individuals can produce the anomalous effect. Nonetheless, consistent positive achievement over large databases does distinguish a few individuals. While the clear differentiation for the selected group of three operators is predetermined by the selection process, their distinctive performance is instructive. Inspection of their effect sizes shows them to fall within the range of the full distribution of operators, indicating that the differential is driven mainly by the large database sizes.
5. The gender variable, necessarily represented in the ANOVA by composite rather than average operator scores, does not show a significant contribution to the variance. Specialized analyses addressing the *average* effect size reveal that there are gender differences, however, comprising small variability and regular correlation with intention for men, and

larger variability and less consistent correlations for women (Dunne, 1998).

6. The temporal development of effect sizes shows a consistent pattern, with initial significant success that declines but then recovers. A specialized analysis reveals that this is a broadly distributed pattern in the intentional conditions, which does not appear in baseline data (Dunne *et al.*, 1994).
7. The overall findings show essentially no evidence for a dependence on spatial or temporal separation, complementing other indications that ordinary physical variables have little impact on these anomalous interactions. A specialized analysis has supplemented this conclusion within REG and other databases (Dunne and Jahn, 1992).

In summary, we conclude that both the comprehensive ANOVA technique and the more sharply focused *ad hoc* analyses can play important roles in the assessment of complicated databases of this sort. On the one hand, analysis of variance allows an examination of large and complex data sets as a whole, with a perspective that displays the relative strength of effects from all the measured variables in the experiment. But ANOVA has limitations, especially in a database where the cell populations in the analytical matrix vary as much as do those in the REG ANOVA. The practical requirements of the experimental program have dictated an *ad lib* accumulation of data under various combinations of parameters in order to explore the engineering questions motivating the experiment, while also maintaining a viable psychological context for the human operators. In the course of the experimental development, emphases have changed to include a progressively more incisive examination of subjective factors. Thus, the REG experiment, although based on an unchanging fundamental design, has a complex and unbalanced, non-orthogonal set of variables. This prevents easy assessment of higher-order interactions, even when it is apparent that interactions among the various conditions may be important contributors to the explanation of variance.

It thus follows that specialized, detailed analyses can help reveal the structure of the data in delimited subsets of the database, without compromising the integrity of interpretations. For example, the implications of spatial separation are critical to modeling the anomalous effects, and an analysis that examines both the linear and higher-order regression of effect size on distance provides essential information. Likewise, the serial position effects must be important indicators of psychological factors bearing on the results, and these can be detailed only in *ad hoc* formats. In this context, ANOVA not only provides guidance for focused assessments, showing, for example, that a specialized analysis must address the non-deterministic and deterministic sources separately, but it also confirms the legitimacy of the *ad hoc* studies. Thus, in these and other cases previously discussed, we see a complementary balance between the global perspective provided by ANOVA and the sharply focused, incisive answers provided by well-posed supplementary analyses.

Acknowledgments

The Princeton Engineering Anomalies Research program has been supported by a number of foundations and individuals, including the Institut für Grenzgebiete der Psychologie und Psychohygiene, the Lifebridge Foundation, the Ohrstrom Foundation, Mr. Richard Adams, Mr. Laurance Rockefeller, and Mr. Donald Webster. We acknowledge also the contributions of the volunteer operators, and the editorial assistance of Elissa Hoeger.

References

Dunne, B. J. (1991). Co-operator experiments with an REG device. *PEAR Technical Report 91005.* Princeton Engineering Anomalies Research, Princeton, NJ.

Dunne, B. J. (1998). Gender differences in human/machine anomalies. *Journal of Scientific Exploration, 12,* 1, 3–55.

Dunne, B. J., Dobyns, Y. H., Jahn, R. G., & Nelson, R. D. (1994). Series position effects in random event generator experiments, with appendix by Angela Thompson. *Journal of Scientific Exploration, 8,* 2, 197–216.

Dunne, B. J., & Jahn, R. G. (1992). Experiments in remote human/machine interaction. *Journal of Scientific Exploration, 6,* 4, 311–332.

Ibison, M. (1998). Evidence that anomalous statistical influence depends on the details of the random process. *Journal of Scientific Exploration, 12,* 3, 407–424.

Jahn, R. G., Dobyns, Y. H., & Dunne, B. J. (1991). Count population profiles in engineering anomalies experiments. *Journal of Scientific Exploration, 5,* 2, 205–232.

Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H., & Bradish, G. J. (1997). Correlations of random binary sequences with prestated operator intention: A review of a 12-year program. *Journal of Scientific Exploration, 11,* 3, 345–367.

Nelson, R. D., Bradish, G. J., & Dobyns, Y. H. (1989). Random event generator qualification, calibration and analysis. *PEAR Technical Report 89001*. Princeton Engineering Anomalies Research, Princeton, NJ.

Nelson, R. D., Bradish, G. J., & Dobyns, Y. H. (1992). The portable PEAR REG: Hardware and software documentation. *PEAR Internal Document 92-1*. Princeton Engineering Anomalies Research, Princeton, NJ.

Nelson, R. D. (1993). CONTCAL: Continuous automatic calibrations, pseudo-intentions, and active experiments. *PEAR Internal Document 93-4*. Princeton Engineering Anomalies Research, Princeton, NJ.

Nelson, R. D., Jahn, R. G., Dunne, B. J., Dobyns, Y. H., & Bradish, G. J. (1997). FieldREG II: Consciousness field effects, replications and explorations. *Journal of Scientific Exploration, 12,* 3, 425–454.

Nelson, R. D., Dobyns, Y. H., Dunne, B. J., & Jahn, R. G. (1991). Analysis of variance of REG experiments: Operator intention, secondary parameters, database structure. *PEAR Technical Report 91004.* Princeton Engineering Anomalies Research, Princeton, NJ.

Radin, D. I., & Nelson, R. D. (1989). Consciousness-related effects in random physical systems. *Foundations of Physics*, *19,* 12, 1499–1514.