Journal of Scientific Exploration, Vol. 12, No. 1, pp. 3–55, 1998 0892-3310/98

Gender Differences in Human/Machine Anomalies

BRENDA J. DUNNE

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

Abstract — Assessment of 270 individual databases produced by 135 human operators in five local and four remote human/machine anomalies experiments conducted in the PEAR laboratory between 1979 and 1993 reveals several significant gender-related differences in performance. Although the 140 databases produced by 62 females are much larger on average than the 130 produced by 73 males, the male average results display significantly stronger correlations with the operators' pre-recorded intentions to shift the output distribution means of a variety of random devices to higher or lower values. Both groups demonstrate greater success in the high-intention efforts than in the low, but whereas a majority of the males succeed in both directions of effort, producing intentional results that are relatively symmetrical in comparison with their empirical baselines, most of the females' low-intention results are opposite to intention. The baseline data generated by the males largely concur with calibration and theoretical expectations, while the females tend to higher than chance values. The female data also frequently display larger score distribution variances. These disparities are more pronounced in five local experiments than in four remote databases. No gender differences appear in two experiments that yield null overall results, suggesting that the gender-related patterns observed in the successful experiments may be indicative characteristics of the primary human/machine anomalies.

Keywords: gender — human/machine interactions — engineering anomlies research

Introduction

The Princeton Engineering Anomalies Research (PEAR) program was established in 1979 to assess the potential vulnerability of sensitive engineering systems and information processors to anomalous influences associated with the consciousness of their human operators. This engineering orientation has focused mainly on the physical parameters of these human/machine interactions, rather than on possible psychological or physiological correlates, other than the primary variable of operator intention. All of these human/machine experiments involve carefully calibrated devices based on well-understood physical processes, each capable of rapidly generating, displaying, and recording extensive sequences of random events. Volunteer human operators attempt, solely through conscious effort, to shift the output distribution means of these devices to higher or lower counts, or to generate an undisturbed baseline, in accordance with pre-recorded intentions, and the data are then examined for statistical correlations between those intentions and device performance. Although these databases are extraordinarily large, consistent with the need for reliable statistical estimates of minuscule effects, they have been produced by a relatively small number of operators. Specifically, nearly 20 million experimental data points, generated between 1979 and 1993 by some 135 operators on a variety of such physical systems, have provided persuasive statistical evidence for small but repeatable shifts of the output distribution means that correlate with the operator intentions.

Previous examination of the individual operator contributions to these databases established that their effects distributed normally around the shifted means, implying that a majority of the operators contributed incrementally to the overall results, in contrast to any dominating performances by a few exceptional operators [1]. While some qualitative indications of characteristic differences in individual performance were noted, particularly among the more prolific operators, these proved difficult to assess quantitatively because of the small signal-to-noise ratios involved. Nonetheless, since the operator pool is fairly evenly composed of 72 males and 62 females, there is an adequate basis for exploring possible collective differences in performance as a function of gender.

The study reported here was also motivated by a body of so-called "co-operator" experiments, wherein pairs of operators addressed the tasks with shared intentions [2]. Beyond providing further confirmation of anomalous correlations between operator intentions and mean shifts, these studies showed no evidence of any simple additive effects of individual operator performance, but did provide strong indications that operator gender may be an important contributing factor. For example, operator pairs of the same sex tended to produce null results, trending insignificantly in the directions opposite to intention. Opposite-sex pairs, on the other hand, produced significant overall results in the desired directions, with effects considerably larger than those generated by these same individuals working alone, and this enhancement of effect size was strongest when the two operators shared a deep emotional bond with each other. Another curiosity of these opposite-sex data was a relative symmetry between the high- and low-going achievements, unlike the asymmetrical yields frequently observed in the single-operator experiments where one intention was typically found to produce considerably stronger results than the other. Prompted by these findings, a comprehensive evaluation of all of PEAR's existing databases has been undertaken to assess the relative performance of its male and female operators over nine different experiments whose design, protocols, and overall results have been detailed previously [3-8].

Methodology

The most direct assessment of male/female differences in performance would appear to be a simple comparison of the composite results of the two groups for each experiment via a simple *z*-score calculation for the differences. However, these composite values are strongly weighted by substantial disparities in the sizes of the individual operator databases, which can easily distort their interpretation. More informative indications of the relative contributions of the male and female operators can be obtained by examining their results on an individual basis and then comparing the average yields and the proportions of operators in each gender group who produce results correlating with intention. This proportional approach also permits comparisons across diverse databases where calculations of effects are necessarily based on different scales.

In the sections to follow, the results of each of nine distinct experiments are presented by gender, both in terms of their composite and average results, and as summaries of the proportional yields of the individual operators. (Full details of the individual results are available in a Technical Report [10]). It should be noted at the outset that most of the experimental databases are relatively small in terms of the numbers of contributing operators, and thus the statistical results based on these proportions frequently entail large error bars. It should also be noted that many of the operators participated in more than one experiment, but since all of the experiments are independent of each other, each operator-experiment database is treated as a separate entity. This approach results in a total of 270 individual contributions over nine separate experiments, 130 from male operators and 140 from female, comprising a more robust base for overall statistical assessment of gender contributions.

1. Random Event Generator Experiments

The most extensive PEAR databases have utilized a microelectronic random event generator (REG) as the target device [3-6]. The "benchmark" experiment comprises more than 2.5 million trials, each consisting of 200 random binary samples. These data were generated over a 12-year period by 91 operators in 522 independent experimental series ranging in size from 1000 to 5000 trials per intention, depending on the protocol involved. (In all PEAR experiments, a "series" is the pre-established evaluative unit, each constituting an independent replication of the basic experiment.) The benchmark REG database was accumulated over three distinct experimental phases which differed in terms of series size, run length (the number of trials produced automatically as a result of a single initiating button push), and the number of secondary options available to the operator, *e.g.,* run length, automatic or manual operation, volitional or instructed assignment of intention, or the available modes of visual feedback on the machine face and its accompanying computer screen. However, all the experiments followed the same basic tri-polar protocol in which the operator was seated in the same room as the device and generated data under three distinct intentions: attempts to shift the mean of the output distributions in the positive direction (HI), in the negative direction (LO), or to produce a baseline (BL) under no directional intention, with all other conditions held constant for the duration of a given series.

Of the 91 operators who contributed to this database, 50 males produced a total of 228 series, or approximately 327,000 trials per intention, and 41 females generated 294 series, or approximately 506,000 trials per intention. (These numbers are approximate because in some of the earlier series a randomly assigned instruction for the direction of each run resulted in unequal numbers of trials per intention; a later modification of the program guaranteed equal numbers of trials per intention in this "Instructed" mode.) The composite results of this database, as well as the relative contributions by male and female operators, are summarized in Table 1. The "normalized deviation," δ_c, utilized here is simply the deviation of the composite experimental mean from the theoretical expectation of 100, multiplied by 100 for convenience of tabulation. It provides an indication of the magnitude of the deviation achieved, but is vulnerable to statistical uncertainty for small data sets.[1] The "z-score", or z_c, defined as the deviation of the composite experimental mean from the theoretical expectation normalized by the standard error, , where σ_0 is the theoretical trial standard deviation and N is the number of trials in the given data set, provides a more reliable indication of the statistical significance of the achieved deviation over databases of varying sizes but, as noted above, can obscure the absolute magnitude of achievement in the smaller data sets. These two indicators, δ_c and z_c, thus complement one another for interpretation of the results.

Table Notes. In this and subsequent tables, achievements in the direction of effort in the high intentions (HI) and in the high-low differences (HI-LO) are indicated by positive deviations and z-scores, and in the low intentions (LO) by negative numbers. Positive numbers in the baselines (BL) indicate results higher than the theoretical mean. Results opposite to intention, or lower than the theoretical expectation in the baselines, are indicated by parentheses. Those z-scores exceeding the one-tailed $p<.05$ criterion (>±1.65) in the direction of intention, and their associated probabilities, are noted by asterisks (*); z-scores >±1.65 opposite to intention are indicated by daggers (†). Probabilities for intentional efforts are calculated on a one-tailed basis; those for baselines, where there is no directional expectation, are two-tailed, with a $p<.05$ criterion of $z>\pm 1.96$.

These composite results suggest that while both groups produce comparable results in the LO and BL, the female operators are collectively more suc-

[1]This normalized deviation is similar to the standardized "effect size" defined by Rosenthal [11], except that his version is normalized by the theoretical trial standard deviation, σ_0, while ours is normalized by an arbitrary constant for convenience in tabulation. Since σ_0 is itself a constant of the experiment, the normalized deviations, δ_c, and standard effect sizes, ε, are related by the constant ratio of $\delta_c/\varepsilon = 100\sigma_0 = 707.1$.

TABLE 1
Composite Results of All Local REG Experiments

	HI	BL	LO	HI–LO
All Operators (522 Series)				
Number of Trials (N)	839,800	820,750	836,650	~837,825
Distribution Mean (μ)	100.026	100.013	99.984	.042
Normalized Deviation (δ_c)	2.6	1.3	–1.6	4.2
Std. Dev. Trial Scores (σ)	7.070	7.074	7.069	9.998
z-Score (z_c)	3.369	1.713	–2.016*	3.809*
Probability (p)	4×10^{-4}*	0.086	0.022*	7×10^{-5}*
50 Male Operators (228 Series)				
Number of Trials (N)	331,650	316,750	331,300	~331,475
Distribution Mean (μ)	100.015	100.011	99.983	0.032
Normalized Deviation (δ_c)	1.5	1.1	–1.7	3.2
Std. Dev./Trial Scores (σ)	7.060	7.064	7.064	9.987
z-Score (z_c)	1.228	0.865	–1.424	1.875*
Probability (p)	0.110	0.386	0.077	0.030*
41 Female Operators (294 Series)				
Number of Trials (N)	508,150	504,000	505,350	~506,750
Distribution Mean (μ)	100.033	100.015	99.986	0.047
Normalized Deviation (δ_c)	3.3	1.5	–1.4	4.7
Std. Dev./Trial Scores (σ)	7.077	7.080	7.073	10.006
z-Score (z_c)	3.339*	1.500	–1.441	3.382*
Probability (p)	4×10^{-4}*	0.134	0.075	4×10^{-4}*

* — see Table Notes on p. 6.

cessful than the males in the HI efforts, resulting in a corresponding advantage in the HI–LO. However, as noted above, this impression is misleading because of the considerable variability among individual operator performances and in the sizes of their respective databases. The female average database is nearly twice as large as the male average and includes three exceptionally large individual databases with strong positive results. Even excluding the most prolific female database consisting of some 120,000 trials per intention, the average female database still remains nearly a third larger than the average male's. While this difference clearly cannot be regarded as an experimental result, it bears noting because of its impact on the statistical representation of the composite results; it may also reflect different operational strategies employed by the two groups.

The individual operator performances are summarized by gender in Table 2, wherein are displayed the averages of the individual normalized deviations and the average z-scores for the HI, BL, and LO efforts, along with those of the HI–LO differences for each group. The number and proportion of operators of each gender who produce results consistent with their intentions (or above 100 in the baselines), relative to the 50% who might be expected to do so by chance, and the number and proportion of individuals who achieve results beyond the one-tailed .05 chance expectation (two-tailed for baselines) are also provided, with the proportions in the opposite tail in parentheses, along with

TABLE 2
Individual Operator Results of All Local REG Experiments, by Gender

	50 Male Operators			
	HI	BL	LO	HI-LO
Average *N*	6,594	6,373	6,607	~6,600
Average δ	3.6	2.1	–0.9	4.5
Average *z*	0.27	0.18	–0.13	0.28
# Oprs p<0.50	34	26	29	33
Proportion	0.68	0.52	0.58	0.66
Prop. *z*	2.55*	0.28	1.13	2.26*
# Oprs p<0.05	7 (3)	3 (0)	0 (1)	3 (0)
Proportion	0.14 (0.06)	0.06 (0.00)	0.00 (0.02)	0.06 (0.00)
Prop. *z*	2.43*(0.32)	0.32 (–2.23)†	–2.23 (–1.11)	0.32 (–2.23)†
	41 Female Operators			
	HI	BL	LO	HI-LO
Average *N*	12,384	12,293	12,335	~12,360
Average δ	0.4	4.2	(3.4)	(–3.1)
Average *z*	0.20	0.35	(0.16)	0.02
# Oprs p<0.50	23	27	14	14
Proportion	0.56	0.66	0.34	0.34
Prop. *z*	0.77	2.05†	(–2.05)†	(–2.05)†
# Oprs p<0.05	4 (1)	0 (0)	3 (3)	3 (3)
Proportion	0.10 (.02)	0.00 (.00)	0.07 (.07)	0.07 (0.07)
Prop. *z*	1.25 (–0.84)	–2.00† (–2.00)†	0.64 (0.64)	0.64 (0.64)
M/F Diffs.				
z_{diff} p<0.50	1.14	1.33	2.28*	3.04*
z_{diff} p<0.05	0.96 (0.80)	1.58 (–0.31)	–2.08*(–1.25)	–0.19(–2.08)†

* and † — see Table Notes on p. 6.

the statistical *z*-scores associated with those proportions.[2] The statistical merit of these proportional gender differences are displayed as *z*-scores (z_{diff}) at the bottom of the table. For the p<.50 criterion, these are determined by comparing the proportion of operators in each group who succeed in the direction of effort and calculating a *z*-score for the proportional difference:

These individual operator summaries treat each operator's database, regardless of size, as a single contribution to its respective gender group, thus eliminating the disproportionate contributions of the more prolific operators to the

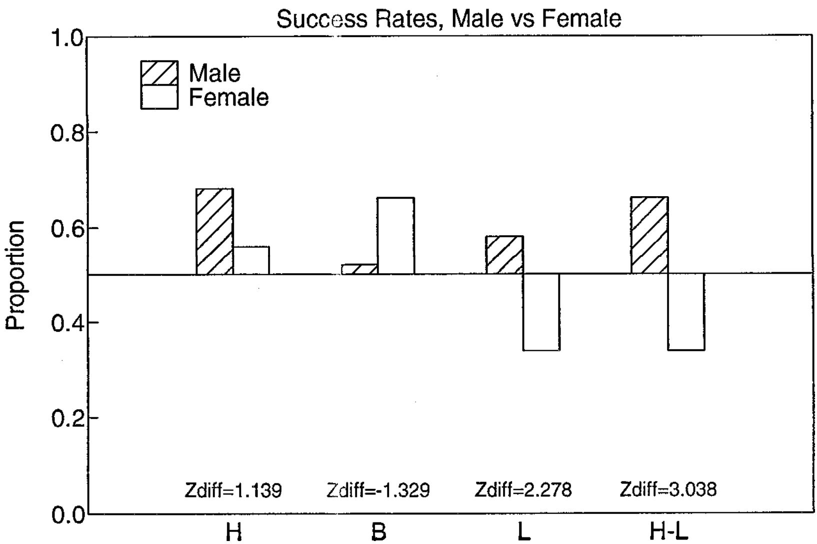
[2]Because the binomial distribution for these extreme-tail populations is seriously distorted from the normal approximation, these *z*-scores are not calculated in the usual way, but instead are back-computed from the exact binomial probability of the number of observed "successes" (operators with p<.05), with a further correction for discreteness. The z_{diff} entries for p<.05 are likewise obtained by comparison of these back-calculated *z*-scores, rather than from a binomial approximation.

composite results, and provide quite a different picture of gender performance. In contrast to the impression of stronger female performance suggested by the composite comparisons, the average male operator actually proves to be more successful than the average female in producing results consistent with intention, and in generating baselines consistent with theoretical expectation. For example, the average male δ_H of 3.6 is almost an order of magnitude larger than the female δ_H of 0.4, although this difference is not statistically significant. In the LO efforts, the average male achieves a modest δ_L –0.9 in the direction of effort, while the average female δ_L of 3.4 is opposite to intention, resulting in a significant z_{diff} of 2.28. These contrasts carry over into comparisons of the $\delta_{H\text{-}L}$ where the male average is 4.5 in the direction of effort, and the female average is –3.1, opposite to intention, yielding a highly significant z_{diff} of 3.04. The BL's of both groups are above the theoretical value, but the average female δ_B of 4.2 is twice as large as the male 2.1. Although the difference between the groups is not statistically significant, the dissimilarities in the baseline performance of the two groups may be of some interest, given the ostensibly null intention prevailing in this condition. Consistent with chance expectations, only 52% of the males generate baselines above 100, while a significant majority of the females (66%) exceed the theoretical value.

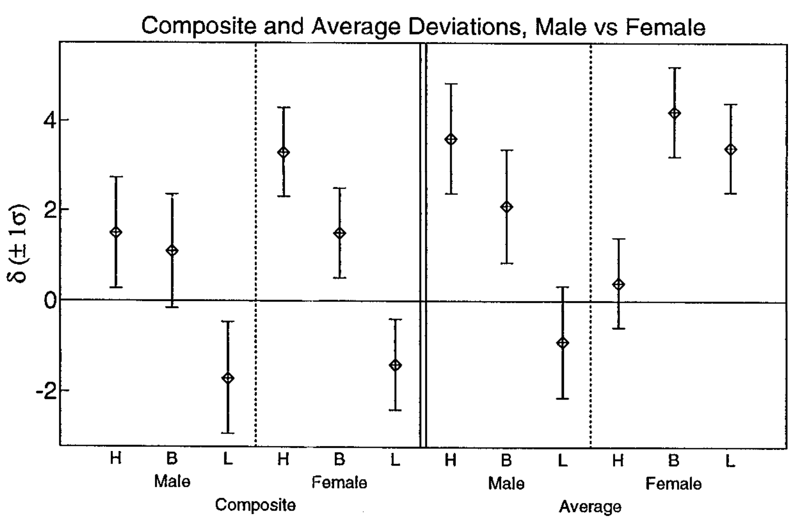
At the $p<0.05$ level the individual achievements and group differences are considerably less robust, and any attempt at interpretation is correspondingly more ambiguous since, given the small size of the two populations, two or three operators in either group would be expected to exceed this criterion by chance in both tails of the population. Although seven males (14%) exceed the $p<0.05$ level of achievement in the HI intention ($z_M = 2.43$), even this result must be interpreted very cautiously given the multiple analyses involved. It is perhaps worth noting, however, that of the full pool of 91 operators, 11 (12%) produce significant results in the HI efforts ($z = 2.65$), while the number of operators of either sex exceeding $p<0.05$ in the LO's is well within chance.

The different performances of the male and female operators are illustrated graphically in Figure 1, where the top portion displays the proportional success rates for both groups, and the bottom portion their relative composite and average deviations with 1σ error bars. Figures 2a and b, 3a and b, and 4a and b, plot the individual male and female operator effect sizes as a function of database size for the HI, LO, and BL. Overlaid on these are the chance mean and empirical mean levels, along with the corresponding .05 tail probability envelopes. The inset tables list the number of operators with effect sizes above/below the two mean values. In these representations, a number of distinctions between male and female performance, both qualitative and quantitative, are clearly evident, especially in the LO and BL.

Previous analyses of this REG database demonstrated a significant series position effect, where strong yields in operators' initial series tended to be followed by declines in the second and third series, with recovery to positive, but more modest yields in subsequent series [6, 12]. Since a substantial majority

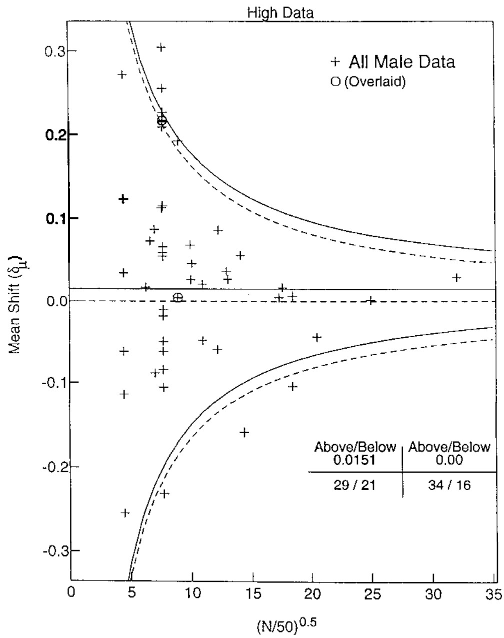


(1a)

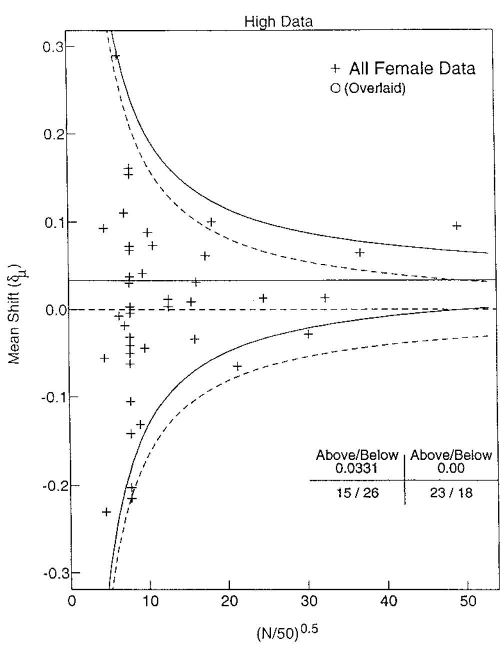


(1b)

Fig. 1. Gender Comparisons in Local REG Results.

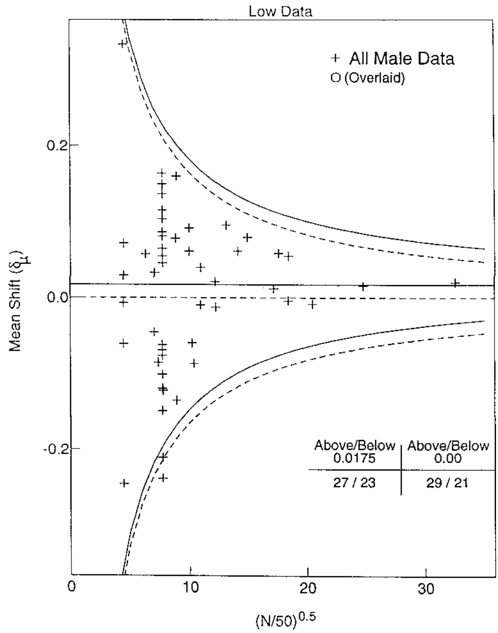


(2a)

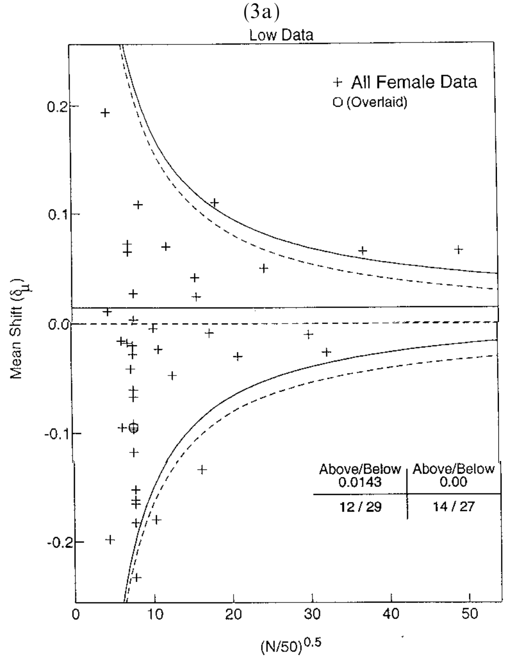


(2b)

Fig. 2. Male and Female Operator Performance: Local REG, High Intention.

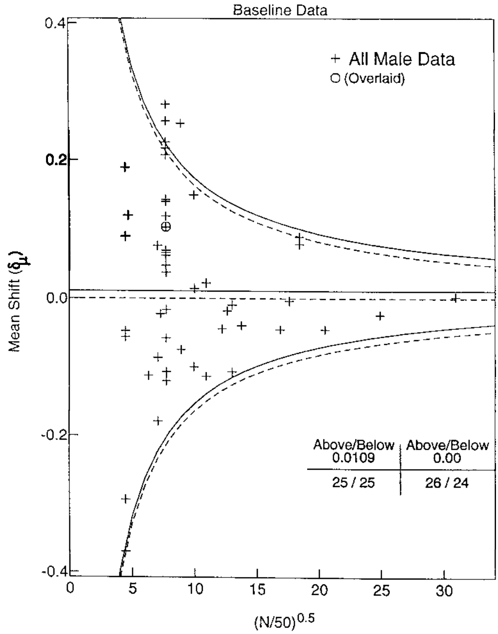


(3a)

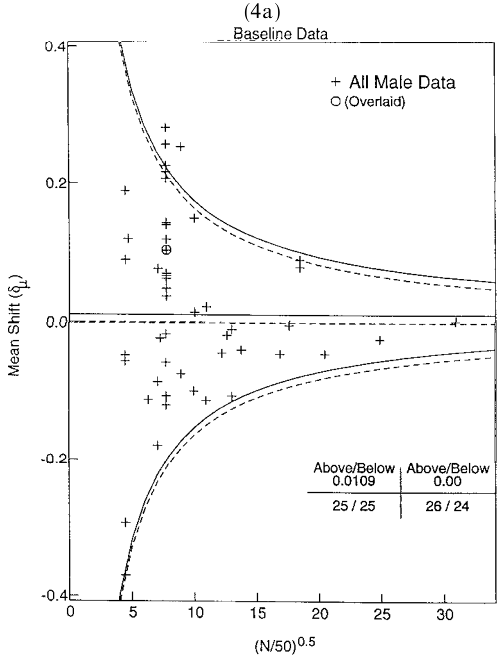


(3b)

Fig. 3. Male and Female Operator Performance: Local REG, Low Intention.



(4a)



(4b)

Fig. 4. Male and Female Operator Performance: Local REG, Baseline.

of both the male and female operators in this pool produced only three or fewer series, it is worth considering whether the apparent gender-related differences could be an artifact of such a series position effect. To address this concern, the normalized deviations and statistical z-scores associated with only the first series produced by each operator were examined separately. Although these initial series also vary somewhat in the numbers of trials involved, the combinations of the δ criterion and the statistically normalized z-score provide reasonable representations of the results for purposes of comparison with those of the full operator databases (Table 3).

These first-series comparisons clearly bear sufficient similarity to those of

TABLE 3
Local REG Results, by Gender: First Series *vs.* Full Databases

50 Male Operators				
	HI	BL	LO	HI–LO
Full Databases				
Composite δ_c	1.5	1.1	–1.7	3.2
Composite z_c	1.23	0.87	–1.42	1.88*
Average δ	3.6	2.1	–0.9	4.5
Average z	0.27	0.18	–0.13	0.28
Prop. Oprs $p<0.50$	0.68*	0.52	0.58	0.66*
Prop. Oprs $p<0.05$	0.14* (0.06)	0.06 (.00)	0.00 (.02)	0.06 (.00)
First Series Only				
Composite δ_c	5.7	2.2	–3.2	8.9
Composite z_c	2.29*	0.81	–1.29	2.53*
Average δ	6.2	2.1	–3.2	9.4
Average z	0.33	0.11	–0.18	0.36
Prop. Oprs $p<0.50$	0.62*	0.63*	0.54	0.66*
Prop. Oprs $p<0.05$	0.08 (0.00)	0.00 (0.00)	0.08 (0.02)	0.06 (0.02)

41 Female Operators				
	HI	BL	LO	LO–BL
Full Databases				
Composite δ_c	3.3	1.5	–1.4	4.7
Composite z_c	3.34*	1.5	–1.44	3.38
Average δ	0.4	4.2	(3.4)	(–3.0)
Average z	0.20	0.34	(0.16)	0.02
Prop. Oprs $p<0.50$	0.56	0.66*	0.34†	0.34†
Prop. Oprs $p<0.05$	0.10 (0.02)	0.00 (0.00)	0.07 (0.07)	0.07 (0.07)
First Series Only				
Composite δ_c	3.6	2.7	1.2	2.4
Composite z_c	1.57	1.13	0.51	0.75
Average δ	2.1	3.9	(2.8)	(–0.7)
Average z	0.18	0.20	(0.13)	0.03
Prop. Oprs $p<0.50$	0.54	0.59	0.41	0.41
Prop. Oprs $p<0.05$	0.07 (0.02)	0.02 (0.00)	0.05 (0.10)	0.07 (0.05)

* and † — see Table Notes on p. 6.

the full databases to confirm the significant gender-related differences noted in the overall results. Even in their initial encounters with the REG, the male operators are more successful than the females in producing results corresponding to intention, particularly in the LO efforts and, correspondingly, in the HI–LO comparisons. Statistical comparisons between the proportions of males and females succeeding in the direction of effort ($p<.50$) produce a marginal z_{diff} of 1.62 in the LO intentions and a significant z_{diff} of 2.38 in the HI–LO differences. There are no significant gender differences in the HI (z_{diff} = 0.76) or BL (z_{diff} = 0.38) comparisons.

Since both groups consistently produce better average correlations with intention in the high efforts than in the low, and both tend to distort the baselines in the high direction, albeit to differing degrees, it is also essential to reconfirm the absence of any technical bias in the performance of the REG device itself. Some 5.8 million calibration trials accumulated on this machine over a period of several years yield an overall mean of 99.998, well within chance expectations (z = –0.826), and slightly *below* the theoretical mean of 100. Thus, the high-going asymmetries in the operator-generated data cannot be attributed to machine bias, but must be related to some factor associated with the human operators, a factor which manifests more strongly in female than in male performance.

It is also worth noting the slight disparities in the trial score standard deviations produced by the two groups, as indicated in the composite summaries of Table 1. While none of the F-ratios comparing these values exceed chance expectations, in all three intentions the female distribution trial variances are slightly larger than those of the males, a trend that will bear watching in later experiments.

In summary, a number of suggestive differences emerge from comparisons of male and female performance in these benchmark REG experiments:

1. On average, the female operators tend to be nearly twice as prolific as the males in data generation. (While not an experimental result, this affects the interpretation of the statistical results, and may eventually prove to be a relevant indicator of differences in the strategies deployed by the two groups.)
2. In both the high- and low-intention efforts, the male average normalized deviations and statistical z-scores are larger and more highly correlated with intention than those of the females.
3. Although both groups are more successful in the high-going efforts than in the low, this asymmetry is much stronger in the female data.
4. Consistent with chance expectations, only 52% of the males produce baselines above the theoretical mean, in contrast with a significant proportion (66%) of the females.
5. While 14% of the males exceed the $p<.05$ criterion in the high-intention efforts, in all the other experimental conditions the proportions of operators producing significant results are within chance expectations.

6. Although none of the differences is independently significant, in all three intentions the females produce larger trial score standard deviations than the males.
7. Examination of the results of the first series produced by each operator in this experiment indicates gender-related differences similar to those seen in their full databases, thus discounting the possibility that these disparities are associated with series position effects or are statistical artifacts of the differences in individual database size.
8. Extensive calibration data show no evidence of any bias in device performance, confirming that the trends observed in the experimental data are associated with the human operators.

While this benchmark REG database comprises the largest number of participating operators of all the PEAR experiments, it is still based on the contributions of a relatively small population. Thus, although the observed gender differences are strongly suggestive, they are far from statistically robust. In order to determine to what degree these gender-related trends are representative, it will be useful to compare them with the yields of other PEAR human/machine databases, even though these are yet smaller in terms of operator contributions. Several of these other experiments involve physical devices that lack a theoretical reference and thus require statistical analyses based on differential comparisons of two empirical distributions. Therefore, before such cross-experiment concatenations can be attempted, it will first be necessary to represent these REG results in a similar format. Table 4 presents the composite and average REG results by gender, comparing the high and low efforts with the empirical baselines generated by each operator, rather than with the theoretical value. These comparisons are illustrated graphically in Figure 5.

It is important to emphasize that in these differential calculations the HI–LO, HI–BL and LO–BL comparisons are no longer statistically independent, leaving the results of the HI–LO comparisons as the primary statistical figures of merit in these analyses, and in those of all the other experiments in this survey. Nonetheless, the HI–BL and LO–BL comparisons can be informative indicators of database asymmetries and, when contrasted with the theoretically-based yields of Table 1 and Figure 1, emphasize how shifts of the putatively "null" baselines can affect the relative proportions of "successful" achievements in the directions of intention. The tendency of both groups, especially the females, to produce baseline means higher than the theoretical value here compounds with the variability among the individual operator baselines to present a considerably different picture of the REG yields than that produced by the theoretical comparisons. For example, the average female δ_{H-L} and δ_{H-B} display extra-chance trends opposite to intention in both comparisons, indicating that the majority of female operators are producing substantially asymmetrical patterns of performance. On the other hand, the relatively

TABLE 4
Individual Operator Results of All Local REG Experiments, by Gender
(Referenced to Empirical Baselines)

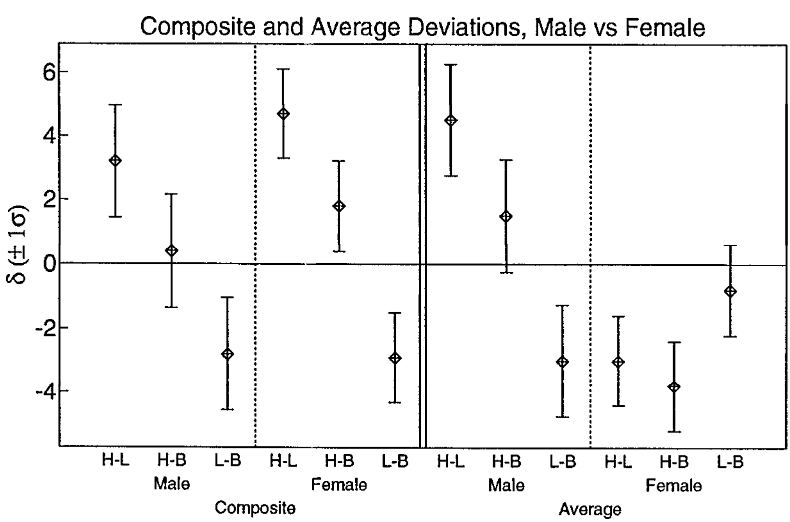
	HI–LO	HI–BL	LO–BL
50 Male Operators			
Avg. *N* per Intention	~6,600	~6,600	~6,600
Composite Diff. (δ_c)	3.2	0.4	–2.8
Composite S.D. (σ_c)	9.987	9.987	9.990
Average δ	4.5	1.5	–3.0
Average z	0.28	0.06	–0.22
# Oprs. p<0.50	33	29	27
Proportion	0.66	0.58	0.54
Proportional z	2.26*	1.13	0.57
# Oprs. p<0.05	3 (0)	0 (1)	8 (0)
Proportion	0.06 (.00)	0.00 (0.02)	0.16 (0.00)
Proportional z	0.32 (–2.23)†	–2.23† (–1.11)	2.88* (–2.23)†
41 Female Operators			
	HI–LO	HI–BL	LO–BL
Avg. *N* per Intention	~12,360	~12,360	~12,360
Composite Diff. (δ_c)	4.7	1.8	–2.9
Composite S.D. (σ_c)	10.006	10.011	10.008
Average δ	(–3.1)	(–3.8)	–0.8
Average z	0.02	(–0.10)	–0.13
# Oprs. p<0.50	14	15	22
Proportion	0.34	0.37	0.54
Proportional z	(–2.05)	(–1.66)†	0.51
# Oprs. p<0.05	3 (3)	2 (3)	4 (0)
Proportion	0.07 (.07)	0.05 (.07)	0.10 (.00)
Proportional z	0.64 (0.64)	–0.03 (0.64)	1.25 (–0.99)
Male/Female Diffs.			
z_{diff} p<0.50	3.04*	1.99*	0.00
z_{diff} p<0.05	–0.19 (–2.08)†	–1.63 (–1.25)	1.30 (–0.99)

* and † — see Table Notes.

symmetrical pattern of average male results exhibits modest but positive correlations with intention in all three differential comparisons.

Specifically, in the differential analyses 58% of the males produce HI–BL results in the direction of effort (z_{H-B} = 1.13), compared to 68% whose HI's exceed the theoretical mean (z_H = 2.55). In the LO–BL, 54% produce separations in the desired direction (z_{L-B} = 0.57), compared to 58% with LO results below the theoretical value (z_L = 1.13). The males HI and LO results thus prove to be even more symmetrical relative to their empirical baselines than to the theoretical mean. On the other hand, only 37% of the females produce HI–BL separations corresponding to the directions of effort (z_{H-B} = –1.66) compared to 56% whose HI's exceed the theoretical mean (z_H = 0.77). In the LO–BL comparisons, 54% are successful (z_{L-B} = 0.51),

(5a)



(5b)

Fig. 5. Gender Comparisons in Local REG Results (Referenced to Empirical Baselines).

compared to only 34% whose LO results are below the theoretical value (z_L = –2.05), thus emphasizing the asymmetry in their intentional performances relative to their empirical baselines. This asymmetry is reflected in the z_{diff}'s of the group proportions, where the male/female difference in the HI-BL yields a z_{diff} = 1.99, but their LO–BL performances are statistically indistinguishable. (Recall that relative to the theoretical mean, the strongest differences between the two groups were in the low-intention efforts, while the high and baseline comparisons were within chance.)

The proportions of significant individual achievements also change with this shift to empirical comparisons, particularly in the male database. Relative to theory, seven males (14%) produce significant δ_H results in the direction of intention, and none in the δ_L. Relative to their respective baselines, however, none of the males achieve significant results in the δ_{H-B}, while eight (16%) produce significant δ_{L-B} separations. By theoretical standards, four females (10%) produce significant δ_H results and three (7%) in the δ_L, while in the empirical comparisons only two (5%) achieve significant $\delta_{H\text{-}B}$ results and four (10%) succeed in the δ_{L-B}. None of the male/female z_{diff}'s are significant.

2. Remote REG Experiments

Another substantial body of data generated on the same REG device consists of 212 experimental series, totaling some 458,000 trials per intention, produced under a "remote" protocol [7]. In the majority of these experiments, comprising 184 series and 396,000 trials per intention, the operators were not present in the laboratory while the machine was in operation, but were directing specific intentions from remote locations for the outcomes of runs generated in the laboratory at pre-arranged times. Some 47 of these series followed an "off-time" protocol where the operators deliberately generated their intentions at times prior to or after machine operation. A hybrid "remote" protocol consisted of an additional 28 series, or 62,000 trials per intention, in which the operators were present in the laboratory complex and personally initiated the REG operation, but were situated in a different room while the device was running. Although these 28 hybrid series were not included in the formal remote database described in Reference [7], they are included in the present survey to extend the sizes of the operator pools. In all of these remote experiments, none of the laboratory staff had knowledge of the operators' intentions until well after the data were produced and recorded.

Of the total of 27 operators contributing to this remote database, 12 males produced a total of 164,000 trials per intention in 72 series, and 15 females a total of 294,000 trials per intention in 140 series. (The earliest remote experiments defined a single series as 3000 runs per intention conducted in three separate sessions, each consisting of 1000 trials per intention generated automatically in three single 1000-trial runs; a later modification defined each such session as an independent series.) The results of these experiments are

TABLE 5
Individual Operator Results of All Remote REG Experiments, by Gender
(Referenced to Empirical Baselines)

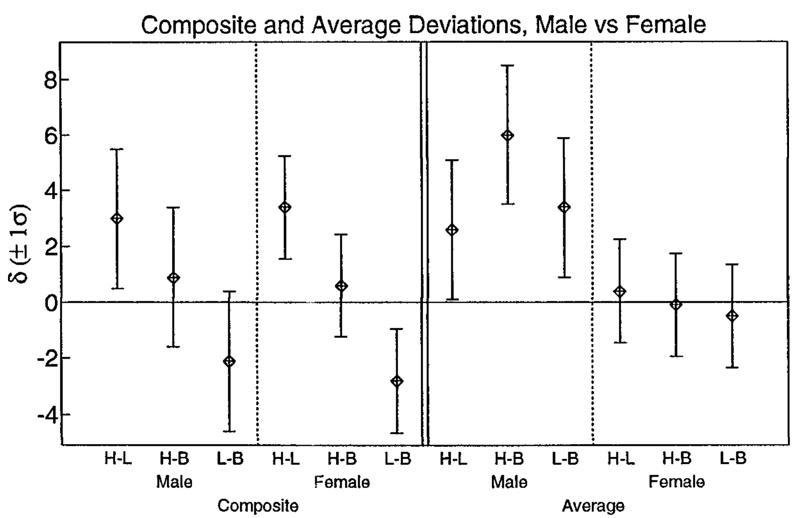
	HI–LO	HI–BL	LO–BL
12 Male Operators			
Avg. N per Intention	13,667	13,667	13,667
Composite Diff. (δ_c)	3.0	0.9	–2.1
Composite S.D. (σ_c)	10.031	10.018	10.015
Average δ	2.7	6.0	(3.4)
Average z	0.37	0.30	–0.06
# Oprs. p<0.50	9	7	7
Proportion	0.75	0.58	0.58
Prop. z-Score	1.73*	0.55	0.55
# Oprs. p<0.05	0 (0)	1 (0)	1 (1)
Proportion	0.00 (0.00)	0.08 (0.00)	0.08 (0.08)
Proportional z	–0.99 (–0.99)	0.48 (–0.99)	0.48 (0.48)
15 Female Operators			
	HI–LO	HI–BL	LO–BL
Avg. N per Intention	19,600	19,600	19,600
Composite Diff. (δ_c)	3.4	0.6	–2.8
Composite S.D. (σ_c)	10.016	10.021	10.025
Average δ	0.4	(–0.1)	–0.5
Average z	0.26	(–0.08)	–0.28
# Oprs. p<0.50	8	7	11
Proportion	0.53	0.47	0.73
Proportional z	0.23	(–0.23)	1.78*
# Oprs. p<0.05	1 (0)	3 (2)	1 (0)
Proportion	0.07 (0.00)	0.20 (0.13)	0.07 (0.00)
Proportional z	0.28 (–1.14)	2.05* (1.24)	0.28 (–1.14)
Male/Female Diffs.			
z_{diff} p <0.50	1.14	–0.62	–0.77
z_{diff} p <0.05	–0.87 (0.19)	–1.21 (–1.58)	0.11 (1.17)

* — see Table Notes on p. 6.

presented in Table 5, in the same differential format as the local REG results in Table 4. Figure 6 displays these comparisons in graphic form.

The small number of operators participating in these remote REG experiments renders any statistical interpretation of the results tentative at best. Nonetheless, the general trends of the two groups bear several potentially relevant similarities to those noted in the local experiments. For example, the average female database is more than 30% larger than the average male's, while the average male δ_{H-L} is 6.75 times larger than that of the average female. A significant proportion of the male operators (75%) produce δ_{H-L}'s in the intended direction, compared with only 53% of the females, although the difference is not significant. Again, these results are quite different from those produced when compared with theory. In the theoretical comparisons, 83% of the

(6a)



(6b)

Fig. 6. Gender Comparisons in Remote REG Results.

males (z_H = 2.29) succeed in the high efforts, but only 50% in the low, while 53% of the females succeed in both the high and low efforts relative to chance. In the empirical comparisons, 73% of them produce δ_{L-B} in the desired direction, compared with only 47% in the δ_{H-B}, while 58% of the males are successful in both comparisons.

It might also be noted that in these remote experiments both groups produce trial score standard deviations larger than the theoretical expectation of 7.071, in this case with those of the males larger than the females' in the HI and LO efforts and that of the females higher in the BL. The male σ_H of 7.095 is significantly larger than chance (p = 0.03), and the σ_L of 7.089 marginally so (p = 0.07). In the baselines, however, the male σ_B of 7.072 is very close to the theoretical value, while the female σ_B of 7.092 is significantly larger than chance (p = 0.02), and considerably larger than those associated with their intentional efforts. (It may be recalled that the female σ_B in the local REG data was also higher than those of their intentional efforts.) None of the F-ratios for the differences between the groups are statistically significant, however.

3. PseudoREG Experiments

In order to address the question of whether the physical behavior of the noise source itself is affected in these anomalous human/machine interactions, the electronic source element was replaced by a categorically different pseudorandom source [3–6]. This device employed a feedback array of 31 microelectronic shift registers that produced a sequence of 2×10^9 bits that cycled continuously with a repetition period of about 60 hours, so that, in principle, the only non-deterministic aspect of the experiment should be the time of incursion initiated by the operator. In its actual operation, however, the ramped sampling mechanism was found to introduce another random element into the process, albeit one of considerably different physical character than the noise diode of the standard REG device. Thus, its label of "pseudorandom" is not technically accurate, but it has proven useful for distinguishing this device from the diode REG and from the fully deterministic ATPseudo experiments described in later sections. Switched into the standard REG apparatus at an appropriate location, this noise source replaced the commercial noise diode and its conditioning circuitry, but left all subsequent sampling, counting, and display circuitry, feedback, and software identical to the benchmark version. From the perspective of the operator, this system was virtually indistinguishable from that of the standard REG, and the experimental protocols employed were identical.

The small database, consisting of 39 series (approximately 102,500 trials per intention), produced on this device by three male and seven female operators, is summarized in Tables 8 and 9, using the same differential analysis employed for Tables 6 and 7.

Given the small male population, the huge error bars make it impossible to calculate meaningful statistics for their tail populations, or to present informa-

TABLE 6
Composite Results of All PseudoREG Experiments

	All Operators	3 Male Oprs	7 Female Oprs
No. Trials/Intention	~102,500	~12,500	~90,000
No. Series	39	5	34
High mean	100.049	100.113	100.040
SD trials	7.049	7.098	7.041
Norm. deviation (δ_c)	4.9	11.3	4.0
z-score	2.25*	1.84*	1.70*
Average δ	11.3	15.8	0.4
Low mean	99.952	100.012	99.944
SD trials	7.074	7.117	7.068
Norm. deviation (δ_c)	–4.8	(1.2)	–5.6
z-score	–2.16*	(0.18)	–2.36*
Average δ	–5.4	(3.6)	–9.2
Baseline mean	99.971	99.946	99.975
SD trials	7.051	7.048	7.051
Norm. deviation (δ_c)	–2.9	–5.4	–2.5
z-score	–1.30	–0.86	–1.07
Average δ	–1.4	–4.2	–0.2
HI–LO δ_c	9.7	10.1	9.6
S.D.	9.986	10.052	9.977
z-score	3.11*	1.17	2.87*
Probability	9×10^{-4}*	0.121	0.002*
HI–BL δ_c	7.8	16.7	6.5
S.D.	9.970	10.003	9.965
z-score	2.51*	1.91*	1.96*
Probability	0.006*	0.028*	0.025*
LO–BL δ_c	–1.9	6.6	–3.1
S.D.	9.988	10.016	9.984
z-score	–0.61	(0.74)	–0.91
Probability	0.271	(0.230)	0.181

* — see Table Notes on p. 6.

tive graphic representation of the results. These are therefore omitted for this experiment and others with similarly small populations. We might simply note the larger size of the average female database and their smaller average deviations in the HI–LO and HI–BL comparisons. In this experiment the female results are more symmetrical than the male relative to their respective baselines, which in both groups are lower than the theoretical value, and their trial score standard deviations are smaller than the theoretical value in all three conditions while those of the males are larger in the HI and LO efforts. All seven of the females and two of the three males produce low-intention results in the direction of effort, and both groups succeed in generating significant composite yields in the high efforts and in the HI–BL comparisons. The composite results of the female low-intention efforts are also significant, as are their composite HI–LO comparisons, and a significant proportion of them produce HI–LO results in the desired direction. A majority in both groups

TABLE 7
Individual Operator Results of All PseudoREG Experiments, by Gender

3 Male Operators			
	HI–LO	HI–BL	LO–BL
Avg # Trials/Int.	4,167	4,167	4,167
Avg Deviation (δ)	22.3	20.0	7.8
Avg z-Score	0.73	1.18	0.46
# Oprs p<0.50	2	2	2
Proportion	0.67	0.67	0.67
Prop. z-Score	‡	‡	‡
# Oprs p<0.05	1 (0)	1 (0)	1 (0)
Proportion	0.33 (0.00)	0.33 (0.00)	0.33 (0.00)
Prop. z-Score	‡	‡	‡
7 Female Operators			
	HI–LO	HI–BL	LO–BL
Avg # Trials	12,857	12,857	12,857
Avg Deviation (δ)	9.5	0.2	–8.9
Avg z-Score	0.95	0.44	–0.51
# Oprs p<0.50	6	5	5
Proportion	0.86	0.71	0.71
Prop. z-Score	1.90*	1.11	1.11
# Oprs p<0.05	1 (0)	2 (1)	1 (0)
Proportion	0.14 (.00)	0.29 (0.14)	0.14 (0.00)
Prop. z-Score	0.92 (–0.71)	2.01* (0.92)	0.92 (–0.71)

* — see Table Notes on p. 6.
‡ insufficient data

produce results in the directions of effort in all three comparisons, with no significant differences between them.

4. ATPseudo Experiments

A related PEAR experiment with a more substantial database involves a computer-generated pseudorandom source developed from a commercial randomization algorithm, seeded by a combination of the current time and microsecond timer count between the setup and start keystrokes by the operator. Unlike the PseudoREG experiment described in the previous section, this ATPseudo experiment is fully deterministic in character. (Its nomenclature derives from the fact that these experiments were run on an IBM 286-AT computer.)

This database consists of 482 series, each of 1000 trials per intention, produced by 17 male and 13 female operators, although 54% of the data were produced by only three female operators. A comprehensive regression-based

TABLE 8
Composite Results of All Local ATPseudoREG Experiments

	All Operators	17 Male Oprs.	13 Female Oprs.
No. Trials	396,000	77,000	319,000
No. Series	396	77	319
High mean	100.004	100.011	100.002
SD trials	7.073	7.083	7.068
Normalized dev.	0.4	1.1	0.2
z-Score	0.33	0.41	0.17
Average dev.	2.1	–1.3	2.3
Low mean	100.014	100.038	100.008
SD trials	7.058	7.070	7.055
Normalized dev.	(1.4)	(3.8)	(0.8)
z-Score	(1.25)	(1.49)	(0.66)
Average dev.	(2.9)	(0.9)	(3.4)
Baseline mean	100.007	100.051	99.996
SD trials	7.072	7.081	7.070
Normalized dev.	0.7	5.1	–0.4
z-Score	0.60	2.02*	–0.33
Average dev.	0.8	2.9	0.3
HI–LO δ	–1.0	–2.7	–0.6
S.D. (σ_c)	9.986	10.052	9.977
z-score	(–0.65)	(–0.76)	(–0.35)
Probability	(0.258)	(0.224)	(0.363)
HI–BL δ	–0.3	–4.0	0.6
S.D. (σ_c)	9.970	10.003	9.965
z-score	(–0.19)	(–1.14)	0.35
Probability	(0.425)	(0.127)	0.363
LO–BL δ	0.7	–1.3	1.2
S.D. (σ_c)	9.988	10.016	9.984
z-score	(0.46)	–0.37	(0.70)
Probability	(0.323)	0.356	(0.242)

* — see Table Notes on p. 6.

analysis of variance of all of PEAR's REG-type experiments indicated that the results of this ATPseudo experiment differed significantly from those of all the others in their lack of any demonstrated anomalous effects [3-6]. Nonetheless, they are included here for completeness, and because they provide a valuable opportunity to compare the gender-related performances observed in successful experiments with those yielding null results. The individual operator results are summarized in Tables 8 and 9 and displayed in Figure 7.

The main point to be noted in these results is that none of the patterns observed in the successful REG experiments are evident in these data.

5. Remote ATPseudo Experiments

Similar null results characterize a smaller remote ATPseudo database consisting of 86 series, generated by 3 males and 7 females, summaries of which

TABLE 9
Individual Operator Results of All Local ATPseudo Experiments, by Gender

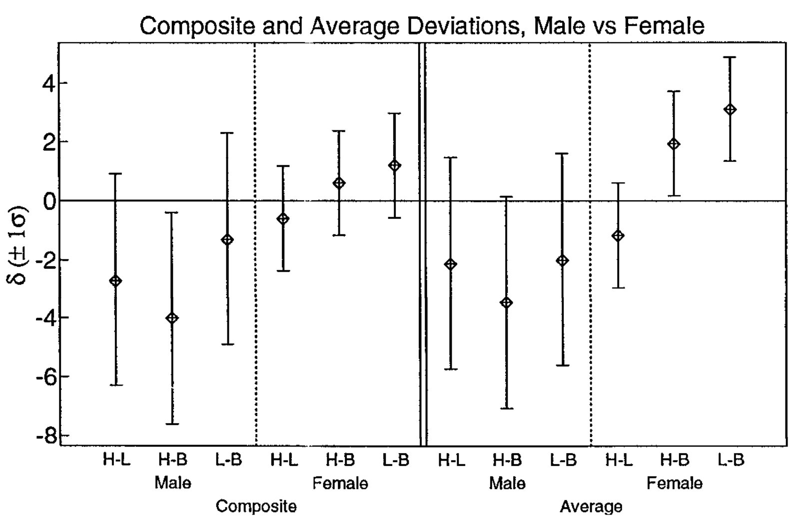
	HI–LO	HI–BL	LO–BL
17 Male Operators			
Avg # Trials	4,529	4,529	4,529
Avg Deviation (δ)	(–2.3)	(–3.5)	–2.0
Avg *z*-Score	(–0.16)	(–0.20)	–0.04
# Oprs p<0.50	8	6	10
Proportion	0.47	0.35	0.59
Prop. *z*-Score	(–0.25)	(–1.24)	0.74
# Oprs p<0.05	0 (0)	0 (0)	1 (0)
Proportion	0.00 (0.00)	0.00 (0.00)	0.06 (0.00)
Prop. *z*-Score	–1.22 (–1.22)	–1.22 (–1.22)	0.16 (–1.22)
13 Female Operators			
Avg # Trials	25,462	25,462	25,462
Avg Deviation (δ)	(–1.2)	1.9	(3.1)
Avg *z*-Score	0.05	0.20	(0.15)
# Oprs p<0.50	5	6	6
Proportion	0.38	0.46	0.46
Prop. *z*-Score	(–0.87)	(–0.29)	(–0.29)
# Oprs p<0.05	0 (0)	1 (0)	0 (1)
Proportion	0.00 (0.00)	0.08 (0.00)	0.00 (0.08)
Prop. *z*-Score	–1.04 (–1.04)	0.41 (–1.04)	–1.04 (0.41)
Male/Female Diffs.			
z_{diff} p<0.50	0.49	–0.60	0.71
z_{diff} p<0.05	–0.23 (–0.23)	–1.19 (–0.23)	0.81 (–1.19)

are presented in Tables 10 and 11. Again, since there are only three male operators, graphic representation is omitted, as are their proportional *z*-scores for $p<.05$.

6. *Random Mechanical Cascade (RMC) Experiments*

To address the relative importance of the physical genre of the particular devices with which the operators attempt to interact, a variety of more diverse machines have been employed, several of which have proven amenable to systematic study and have yielded databases that can be included in this gender effect survey. One of these is a macroscopic random mechanical cascade (RMC) device, measuring some 6′ × 10′ in dimension. This apparatus allows 9000 3/4″ polystyrene spheres to trickle downward through a quincunx array of 330 3/4″ diameter nylon pegs, whereby they are scattered into 19 collecting bins across the bottom, filling them in close approximation to a Gaussian distribution. The growing populations of every bin are tracked photo-electrically and displayed via LED counters at the bottom of those bins, and

(7a)



(7b)

Fig. 7. Gender Comparisons in Local ATPseudo Results.

TABLE 10
Composite Results of All Remote ATPseudoREG Experiments

	All Operators	3 Male Oprs.	7 Female Oprs.
No. Trials	86,000	20,000	66,000
No. Series	86	20	66
High mean	100.027	99.961	100.047
SD trials	7.055	7.037	7.060
Normalized dev.	2.7	(−3.9)	4.7
Average dev.	5.0	(−1.8)	7.0
Low mean	100.015	99.986	100.024
SD trials	7.062	7.041	7.069
Normalized dev.	(1.5)	−1.4	(2.4)
Average dev.	(1.2)	(1.9)	(1.0)
Baseline mean	99.997	99.953	100.010
SD trials	7.071	7.053	7.077
Normalized dev.	−0.3	−4.7	1.0
Average dev.	2.6	−0.4	3.5
HI−LO δ	1.2	−2.5	2.3
S.D. (σ_c)	9.982	9.955	9.991
z-score	0.34	(−0.36)	0.58
Probability	0.37	(0.36)	0.28
HI−BL δ	3.0	0.8	3.7
S.D. (σ_c)	9.989	9.963	9.996
z-score	0.89	0.11	0.95
Probability	0.19	0.46	0.17
LO−BL δ	1.8	3.3	1.4
S.D. (σ_c)	9.994	9.966	10.003
z-score	(0.55)	(0.47)	(0.37)
Probability	(0.29)	(0.32)	(0.36)

simultaneously recorded on-line in an appropriately coded computer file. In the local protocol, the operator is seated on a sofa approximately eight feet from the machine and attempts to distort the distribution of balls to the right or to the left, or to generate a baseline.

The principal RMC database consists of 1131 runs per intention, generated in 87 series by 25 operators, 12 males and 13 females. Each series comprises 20 (in some of the earlier series) or 10 runs per intention [3, 4, 5, 8]. These results have here been combined with those of a smaller, more recent RMC database, consisting of 70 series of only three sets of runs per intention, but following the same basic tri-polar protocol. This extends the operator pools to 16 males and 20 females and offers a slightly stronger statistical base for indications of any gender-related trends.

By its nature, this device precludes any precise theoretical reference. It also displays mild long-term drift in its calibration data, presumably due to mechanical wear, and shorter-term variations correlated with temperature and humidity (which are routinely recorded). As a result, all analytical assessments

TABLE 11
Individual Operator Results of All Remote ATPseudo Experiments, by Gender

3 Male Operators			
	HI–LO	HI–BL	LO–BL
Avg # Trials	6,667	6,667	6,667
Avg Deviation (δ)	(–3.7)	(–1.4)	(2.3)
Avg *z*-Score	(–0.25)	(–0.02)	(0.23)
# Oprs $p<0.50$	1	1	0
Proportion	0.33	0.33	0.00
Prop. *z*-Score	‡	‡	‡
# Oprs $p<0.05$	0 (0)	0 (0)	0 (0)
Proportion	0.00 (.00)	0.00 (0.00)	0.00 (0.00)
Prop. *z*–Score	‡	‡	‡
7 Female Operators			
	RT–LT	RT–BL	LT–BL
Avg # Trials	9,429	9,429	9,429
Avg Deviation (δ)	6.0	3.5	–2.5
Avg *z*-Score	0.33	0.30	–0.03
# Oprs $p<0.50$	4	3	4
Proportion	0.57	0.43	0.57
Prop. *z*-Score	0.37	(–0.37)	0.37
# Oprs $p<0.05$	0 (0)	1 (0)	0 (1)
Proportion	0.00 (0.00)	0.14 (0.00)	0.00 (0.14)
Prop. *z*-Score	–0.71 (–0.71)	0.92 (–0.71)	–0.71 (0.92)

‡ insufficient data

of anomalous effects related to operator intention must proceed on a local differential basis, with only the paired differences among the right (RT), left (LT), and baseline (BL) efforts within a given tri-polar set statistically cumulated and processed. The cumulated differences between the right and left efforts (RT–LT), evaluated in terms of a Student's *t*-test based on the standard deviations of the differences between runs within a local set, are regarded as the primary indicators of any operator effects. However, since each run involves 9000 individual ball trajectories, the *t*-distribution with ~18,000 d.f. in each pair comparison is statistically indistinguishable from the normal *z*-distribution, and so, for consistency of comparison with the REG experiments, the results are represented as *z*-scores in the tables below. Statistical *z*-scores for both the RT–BL and LT–BL are also calculated separately and, while these are not independent of the primary RT–LT comparisons, they can be instructive for evaluating the trends of the three intentions relative to one another.

The composite results of the RMC gender subsets are presented in Table 12 which, for purposes of comparison, also provides the averages of the individual normalized mean shifts for the right, left, and baseline conditions (δ_a), along with their composite values (δ_c). Lacking a theoretical mean value as an

TABLE 12
Composite Results of All Local RMC Experiments

	All Operators	16 Male Oprs.	20 Female Oprs.
No. Run Sets	1341	332	1009
No. Series	157	40	117
Right Mean	10.0180	10.0172	10.0183
SD Run Scores	0.0345	0.0376	0.0337
Composite δ_c	1.8	1.72	1.83
Average δ_a	1.05	1.64	0.86
Left Mean	10.0135	10.0133	10.0136
SD Run Scores	0.0351	0.0310	0.0363
Composite δ_c	1.35	1.33	1.36
Average δ_a	0.64	1.26	0.43
Baseline Mean	10.0184	10.0148	10.0195
SD Run Scores	0.0356	0.0340	0.0364
Composite δ_c	1.84	1.48	1.95
Average δ_a	1.15	1.41	1.07
RT-LT δ_{RL}	0.45	0.38	0.47
S.D. Run Diffs.	0.0492	0.0487	0.0495
z-score	3.348*	1.462	3.022*
Probability	4×10^{-4}*	0.072	0.001*
RT-BL δ_{RB}	(–0.04)	0 0.23	(–0.13)
S.D. Run Diffs.	0.0496	0.0507	0.0496
z-score	(–0.307)	0.922	(–0.882)
Probability	(0.379)	0.178	(0.189)
LT-BL δ_{LB}	–0.48	–0.14	–0.60
S.D. Run Diffs.	0.0500	0.0460	0.0514
z-score	–3.663*	–0.552	–3.907*
Probability	10^{-4}*	0.290	5×10^{-5}*

* — see Table Notes on p. 6.

absolute reference, these deviations are defined as the differences between the experimental means of the bin distributions and the arbitrary value of 10.0, which is essentially the global mean bin of the RMC device, normalized by the latter and expressed in units of 10^{-3} bins/bin. Since the raw run variances are contaminated by environmental drift and machine wear, the standard deviations of the run scores given here are reconstructed from the differential values listed in the lower portion of the table.[3] It should be noted that a *right* intention in this experiment represents an attempt to shift the output distribution mean toward *higher* bin numbers, and a *left* intention toward *lower* bin numbers. Thus, the RT and LT notation in the tables may be regarded as equivalent to the HI and LO indicators of the REG experiments. Table 13 summarizes the individual operator contributions as a function of gender. (Note that again the

[3]Since the spurious contributions to the means cancel out to an excellent statistical approximation within the tri-polar sets, reconstructions of the separate intention variances, σ_h, σ_b, σ_l, from the matrix of the differential variances, σ_{H-L}, σ_{H-B}, σ_{L-B}, are also protected from these artifacts. The appropriate algebraic relations are simply:

$$\sigma^2_h = (\sigma^2_{H-L} + \sigma^2_{H-B} - \sigma^2_{L-B})/2$$
$$\sigma^2_b = (-\sigma^2_{H-L} + \sigma^2_{H-B} + \sigma^2_{L-B})/2$$
$$\sigma^2_l = (\sigma^2_{H-L} - \sigma^2_{H-B} + \sigma^2_{L-B})/2$$

TABLE 13
Individual Operator Results of All Local RMC Experiments,by Gender

16 Male Operators			
	RT–LT	RT–BL	LT–BL
Avg # Run Sets	20.75	20.75	20.75
Avg Difference (δ)	0.38	0.23	–0.15
Avg *z*-Score	0.34	0.22	–0.13
# Oprs $p<0.50$	10	11	10
Proportion	0.63	0.69	0.63
Prop. *z*-Score	1.04	1.52	1.04
# Oprs $p<0.05$	1 (1)	1 (0)	1 (1)
Proportion	0.06 (0.06)	0.06 (0.00)	0.06 (0.06)
Prop. *z*-Score	0.22 (0.22)	0.22 (–1.18)	0.22 (0.22)
20 Female Operators			
	RT–LT	RT–BL	LT–BL
Avg # Run Sets	50.45	50.45	50.45
Avg Difference (δ)	0.43	(–0.21)	–0.64
Avg *z*-Score	0.44	(–0.20)	–0.64
# Oprs $p<0.50$	12	9	14
Proportion	0.60	0.45	0.70
Prop. *z*-Score	0.89	(–0.45)	1.79*
# Oprs $p<0.05$	3 (0)	1 (1)	2 (0)
Proportion	0.15 (0.00)	0.05 (0.05)	0.10 (0.00)
Prop. *z*-Score	1.68*(–1.34)	0.00 (0.00)	0.91 (–1.34)
Male/Female Diffs.			
z_{diff} $p<0.50$	0.18	1.43	0.42
z_{diff} $p<0.05$	–1.11 (1.15)	0.15 (–0.79)	–0.53 (1.15)

* — see Table Notes on p. 6.

difference δ's and their associated *z*-scores displayed here are *averages* of the individual operator differences, and hence are different from the *composite* difference values displayed in the lower half of Table 12.) Figure 7 represents these results in graphical form.

In these RMC experiments, the females are again more prolific than the males in generating data, with average databases some 2.5 times as large, and their intentional results are again less symmetrical relative to their baselines. As in the REG experiments, their stronger composite RT–LT and LT–BL yields and their null RT–BL results give an initial impression of greater success than the males, but again this may well be attributed to their unusual baseline performance. The top portion of Table 12 indicates that both the composite δ_c and average δ_a of the female baselines values are higher than those of either their left- or right-intention efforts, and their composite baseline value is higher than those of any of the three male conditions, strongly suggestive of right-shifted baselines. Although the requisite differential analysis in this experiment does not permit statistical demonstration of such a baseline shift, it may be recalled that a similar pattern was noted in the REG data, where a significant proportion of the female operators produced baseline means higher than the theoretical value. Thus, while the female composite RT–LT and LT–BL separations exceed chance in the intended directions, their RT–BL effects are opposite to intention, resulting in a strong asymmetry in their overall database that bears some resemblance to their REG performance.

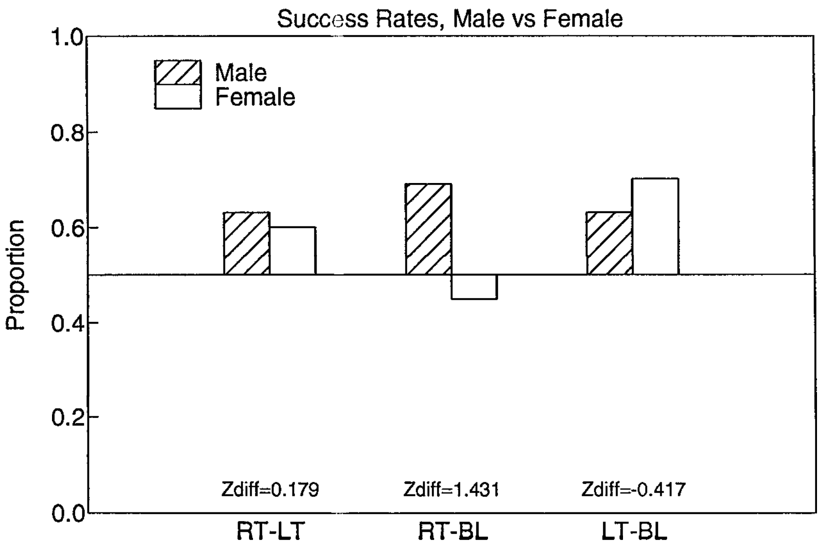
At the level of individual operator performance, a majority of operators in both groups succeed in producing δ_{R-L}'s in the desired direction to a comparable degree, although the male average values in both the RT and LT are substantially larger, *i.e.* more strongly right-shifted, than the female. The RT–BL comparisons indicate that the male average deviation is in the intended direction while the female is opposite to intention, with 69% of the males achieving positive results and only 45% of the females, although the difference is not significant. In contrast, while the majority of operators in both groups are successful in the LT–BL comparisons, the females exceed the males in both the magnitude of their average deviations and in the proportion producing results in the intended direction, although the difference is again statistically negligible.

The reconstructed standard deviations of the female composite score distributions, shown in Table 12, are again larger than those of the males in both the baseline and left-intention efforts. Although the F-ratio of 0.873 (df= 331, 1008) comparing the male and female baseline variances is within chance expectations (p =.136, two-tailed), in the left intention, F = 0.729, with a probability of 6×10^{-4}. In the RT efforts, the male variance is significantly larger than the female (F = 1.245, p = .012).

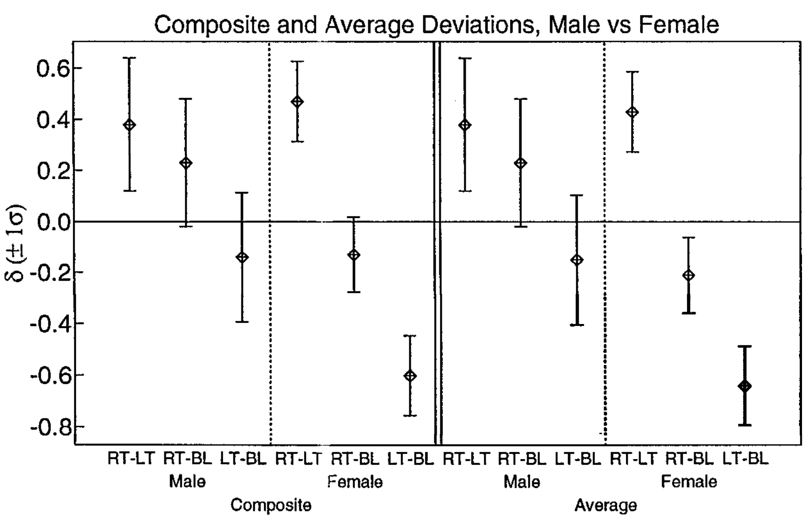
In summary, although the proportional comparisons indicate no significant group differences in this experiment, other than in the variances, the results do suggest some gender-related trends worth noting for their resemblances to, and compounding with, those observed in other experiments. In particular, although a majority of operators in both groups succeed in producing RT–LT separations in the desired directions, comparisons between the composite results of the males' intentional efforts with those of their baselines, shown in Table 12 and Figure 8, indicate a relatively symmetrical pattern of results, while the females again show an apparent preference for shifting the distribution means, especially those of the baselines, toward higher values, thereby producing strongly asymmetrical patterns among their three intentions. Finally, the significant differences in the distribution variances produced by the two groups add further gender-related distinctions in performance.

7. Remote RMC Experiments

A much smaller remote RMC database of 56 series was generated by 11 female and 3 male operators, comprising a total of 337 sets of runs, 285 of which were generated by the females and 52 by the males. The composite results of these experiments are summarized in Table 14 and the relative contributions



(8a)



(8b)

Fig. 8. Gender Comparisons in Local RMC Results.

TABLE 14
Composite Results of All Remote RMC Experiments

	All Operators	3 Male Oprs.	11 Female Oprs.
No. Run Sets	337	52	285
No. Series	56	6	50
Right mean	10.0046	10.0159	10.0025
SD run scores	0.0356	0.0352	0.0394
Composite Dev. (δ_c)	0.46	1.59	0.25
Average Dev. (δ_a)	0.04	1.55	–0.24
Left mean	10.0028	10.0103	10.0014
SD run scores	0.0276	0.0256	0.0326
Composite Dev. (δ_c)	0.28	1.03	0.14
Average Dev. (δ_a)	0.38	1.19	0.11
Baseline mean	10.0022	10.0082	10.0011
SD run scores	0.0394	0.0370	0.0362
Composite Dev. (δ_c)	0.22	0.82	0.11
Average Dev. (δ_a)	0.27	0.99	0.14
RT–LT δ_{RL}	0.19	0.55	0.12
S.D. Diffs.	.0450	0.0435	0.0511
z-score	0.702	0.814	0.416
Probability	0.241	0.208	0.339
RT–BL δ_{RB}	0.25	0.76	0.15
S.D. Diffs.	0.0531	0.0511	0.0535
z-score	0.932	1.122	0.534
Probability	0.176	0.131	0.297
LT–BL δ_{LB}	0.07	0.21	0.03
S.D. Diffs.	0.0481	0.0450	0.0487
z-score	(0.246)	(0.312)	(0.134)
Probability	(0.403)	(0.378)	(0.446)

by gender in Table 15. Again, the small number of participating male operators precludes calculating z-scores for their $p<.05$ proportions or graphical representation.

Despite the small numbers of participating operators in this remote database, like the other small databases they are reported for completeness and for inclusion in the overall concatenations to follow. Once again it may be noted that the males produce larger average deviations than the females in all three comparisons, and the females produce larger databases and larger run standard deviations in the two intentional conditions. While the F-ratio of 0.798 comparing the RT efforts of the two groups is non-significant, in the LT comparisons F = 0.617 (df= 51,284), $p = 0.038$.

8. Pendulum Damping Experiments

Another large database that displays particularly striking gender-related differences has been obtained on a linear pendulum apparatus, constructed for the purpose of determining whether operator intention is capable of influencing its damping rate [9]. The pendulum bob is a 2-inch crystal sphere suspended on a

TABLE 15
Individual Operator Results of All Remote RMC Experiments, by Gender

	RT–LT	RT–BL	LT–BL
3 Male Operators			
Avg # Run Sets	17.33	17.33	17.33
Avg Difference (δ)	0.37	0.57	(0.20)
Avg *z*-Score	0.39	0.57	(0.18)
# Oprs $p<0.50$	2	2	0
Proportion	0.67	0.67	0.00
Prop. *z*-Score	‡	‡	‡
# Oprs $p<0.05$	0 (0)	0 (0)	0 (0)
Proportion	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Prop. *z*-Score	‡	‡	‡
11 Female Operators			
Avg # Run Sets	25.91	25.91	25.91
Avg Difference (δ)	(–0.35)	(–0.38)	–0.03
Avg *z*-Score	(–0.05)	(–0.03)	–0.02
# Oprs $p<0.50$	5	6	6
Proportion	0.45	0.55	0.55
Prop. *z*-Score	(–0.33)	0.33	0.33
# Oprs $p<0.05$	0 (0)	0 (1)	0 (0)
Proportion	0.00 (0.00)	0.00 (0.09)	0.00 (0.00)
Prop. *z*-Score	–0.95 (–0.95)	–0.95 (0.56)	–0.95 (–0.95)

‡ insufficient data

fused silica rod from precision pivots, all enclosed within a clear acrylic box. A high-speed binary counter registers interruptions of photo-diode beams to measure velocities at the nadir of the pendulum arc with microsecond accuracy. The tri-polar protocol requires the operator to alternate attempts to keep swings high, *i.e.* to decrease the damping rate, with attempts to reduce the swing amplitude, *i.e.* to increase the damping rate, relative to undisturbed baseline runs. Data are accumulated in three-minute runs of 100 full swings, and on-line comparisons of the progress of high or low runs with initial baseline runs are processed to provide real-time feedback to the operator in the form of a change in color of the crystal bob. Since the pendulum's performance is highly dependent on local atmospheric conditions, real-time readings of temperature, barometric pressure, and humidity are recorded on-line and data analyses incorporate appropriate adjustments.

Forty operators, 20 males and 20 females, contributed 306 and 609 sets of runs, respectively, for a total database of 915 sets in the local version of this experiment, consisting of 235 complete and five partial series. As originally defined, an experimental series required nine tri-polar sets of 5-minute runs, typically generated in three sessions of three sets each, with a session lasting about 45 minutes. A later modification to the protocol reduced series size to five sets of runs that could be generated in a single session of about 1.25 hours. The five-set series comprise approximately 75% of the database. (The five partial series each consisted of a minimum of five completed sets produced in a nine-set series.)

As in the RMC experiment, the planned analysis compares damping rates of the high and low efforts using paired *t*-tests based on the variance of the differences within sets. For consistency of representation, the resultant *t*-scores have been converted to *z*-scores, using an inverse normal distribution to calculate the equivalent *z*'s corresponding to the *t*-score probabilities.

The composite results of the local pendulum experiments are presented in Table 16 and the individual operator contributions are summarized in Table 17 and displayed in Figure 9. It should be emphasized that the distribution means in these experiments indicate the average damping rates in terms of the *loss* in nadir velocity over the course of the runs. That is, since a "high" intention constitutes an attempt to *decrease* the damping rate and a "low" to *increase* it, success in the high direction produces a larger *negative* number, and vice versa. For convenience of representation and consistency with the other experiments presented in this report, after their initial presentation as negative numbers in the composite means of Table 19, the minus signs are subsequently omitted in the tables and a sign convention employed wherein a deviation in the direction of effort in the HI–LO and HI–BL is indicated by a positive number and *z*-score, and a deviation in the direction of effort in the LO–BL by a negative number and *z*-score. The composite normalized deviations are presented as the actual means minus the nearest arbitrary round value of 40000, multiplied by 10^{-3}, and the run standard deviations are also multiplied by 10^{-3}. The average normalized deviation again refers to the unweighted average value achieved by the operator group and the standard deviations of the run scores are reconstructed from the uncontaminated differential variances, as described in Note 3.

Once again, the females generate much larger databases while the males produce results that better correlate with intention. In this experiment, however, both groups produce lower values in the baselines than in their high or low efforts (recall that the signs are reversed), resulting in strong asymmetries in the performances of both groups, albeit in opposite directions. As in the local RMC experiment, the female composite and average δ's in all three individual intentions are lower than the male, with the lowest in their baselines. They also produce substantially larger standard deviations. In the individual operator databases, the average female differences are negative in all three comparisons, while the average male results are all positive.

A majority of the males (60%) produce HI–LO results in the desired direction, compared to only 30% of the females, yielding a significant z_{diff}, and these disparities are again more pronounced in the HI–BL comparisons. In the LO–BL, only 50% of the males and 40% of the females succeed in the intend-

TABLE 16
Composite Results of All Local Pendulum Experiments

	All Operators	20 Male Oprs.	20 Female Oprs.
No. Run Sets	915	306	609
No. Series	183	61	121
HI Mean	–42236.52	–42164.87	–42272.52
Composite δ_c	2.237	2.165	2.273
Run Score σ	.0699	.0544	.0757
Average δ_a	2.140	2.046	2.233
LO Mean	–42237.82	–42175.40	–42269.18
Composite δ_c	2.238	2.175	2.269
Run Score σ	0.0729	0.0692	0.0742
Average δ_a	2.140	2.058	2.221
BL Mean	–42241.01	–42176.53	–42273.41
Composite δ_c	2.241	2.177	2.273
Run Score σ	0.0783	0.0666	0.0839
Average δ_a	2.142	2.063	2.220
Comp. HI–LO (δ_{HL})	0.001	0.010	–0.004
SD Diffs.	0.101	0.088	0.106
z-score	0.388	2.118*	(–0.799)
Prob.	0.349	0.017*	(0.212)
Comp. HI–BL (δ_{HB})	0.004	0.012	0.000
SD Diffs.	0.105	0.086	0.113
z-score	1.291	2.341*	0.208
Prob.	0.098	0.010*	0.418
Comp. LO–BL (δ_{LB})	0.003	0.002	0.004
SD Diffs.	0.107	0.096	0.112
z-score	(0.902)	(0.142)	(0.967)
Prob.	(0.184)	(0.444)	(0.167)

* — see Table Notes on p. 6.

ed direction, not surprising given the low baseline values produced by both groups.

By the $p<.05$ criterion, fully 25% of the males produce significant separations in the HI–LO separations while none of the females exceed the chance value, resulting in a strongly significant difference between the two groups (p = .001). Again, this effect is driven by the $\delta_{H\text{–}B}$, where fully 35% of the males and none of the females exceed the chance criterion ($p = 5 \times 10^{-5}$). The tendency of both groups to generate low-going baselines in this experiment indicates a curious reversal from the trends observed in the other experiments, tempting speculation that this might be associated with the ambiguity of the experimental task, with its “high” instruction to *decrease* the damping rate and its “low” to *increase* it.

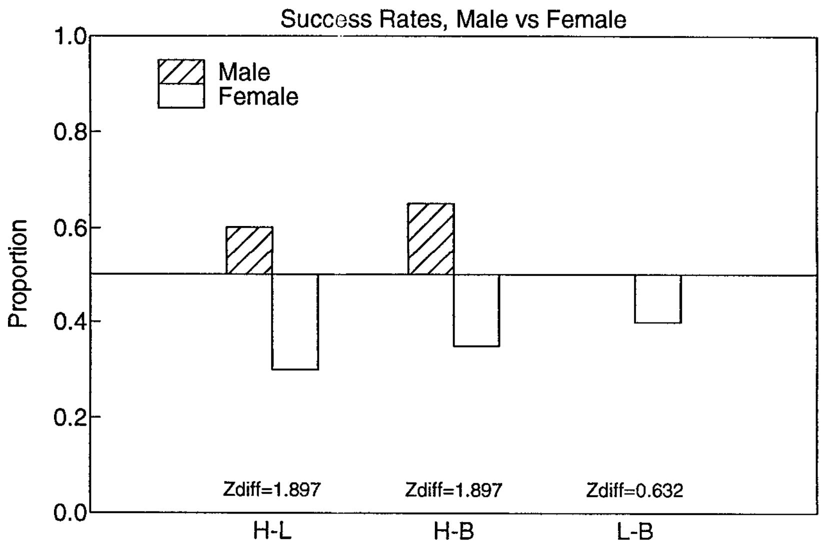
The differential run score standard deviations, shown in Table 16, exhibit extreme discrepancies in this experiment. The $\sigma_M = 0.088$ and $\sigma_F = 0.106$ yield F = 0.689 (df = 305,608), $p = 10^{-4}$, in the primary HI–LO comparison. This contrast is driven both by the HI–BL differences where σ_M = 0.086 and $\sigma_F = 0.113$ (F = 0.579, $p = 6 \times 10^{-8}$), and by the LO–BL where σ_M = .096 and

TABLE 17
Individual Operator Results of All Local Pendulum Experiments, by Gender

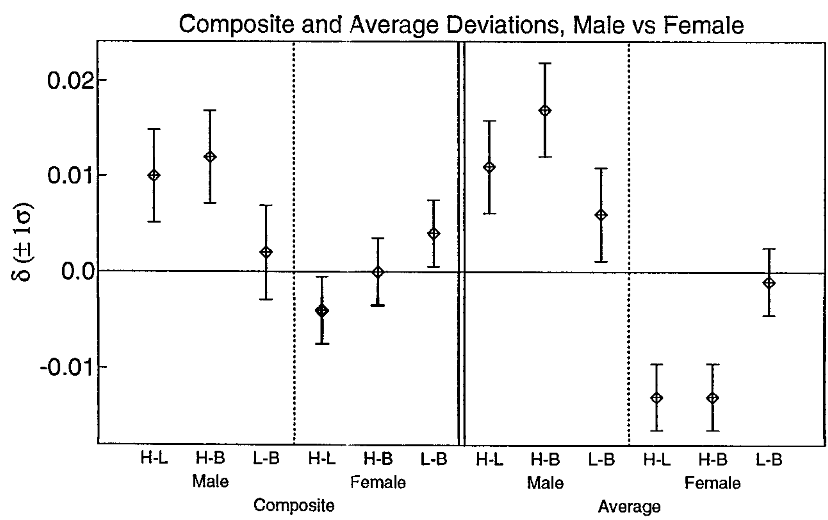
	HI–LO	HI–BL	LO–BL
20 Male Operators			
Avg # Run Sets	15.30	15.30	15.30
Avg Difference (δ)	0.011	0.017	(0.006)
Avg *z*-Score	0.42	0.64	(0.27)
# Oprs $p<0.50$	12	13	10
Proportion	0.60	0.65	0.50
Prop. *z*-Score	0.89	1.39	0.00
# Oprs $p<0.05$	5 (2)	7 (0)	2 (2)
Proportion	0.25 (0.10)	0.35 (0.00)	0.10 (0.10)
Prop. *z*-Score	3.00* (0.91)	4.17* (–1.34)	0.91 (0.91)
20 Female Operators			
Avg # Run Sets	30.45	30.45	30.45
Avg Difference (δ)	(–.013)	(–.013)	–.001
Avg *z*-Score	(–0.46)	(–0.24)	(0.15)
# Oprs $p<0.50$	6	7	8
Proportion	0.30	0.35	0.40
Prop. *z*-Score	(–1.79)†	(–1.34)	(–0.89)
# Oprs $p<0.05$	0 (2)	0 (0)	0 (3)
Proportion	0.00 (0.10)	0.00 (0.00)	0.00 (0.15)
Prop. *z*-Score	–1.34 (0.91)	–1.34 (–1.34)	–1.34 (1.68)*
Male/Female Diffs.			
z_{diff} $p<0.50$	1.90*	1.90*	0.63
z_{diff} $p<0.05$	3.07* (0.00)	3.90* (0.00)	1.59 (–0.54)

* and † — see Table Notes on p. 6.

$\sigma_F = 0.112$ (F = 0.735, $p = 0.001$). When these differential standard deviations are converted to σ's for the individual intentions, following the procedure described earlier, the gender contrast is even more dramatic, with F = 0.516 ($p = 2 \times 10^{-10}$, two-tailed) in the high efforts and F = 0.630 ($p = 6 \times 10^{-6}$) in the baselines. The male/female difference in the low-intention σ's yields a non-significant F of 0.870. In all cases, the female σ's are larger than those of the males and their baseline standard deviations are larger than those of their intentional efforts, consistent with the more modest trends observed in the REG and RMC experiments. Given the huge differences in this experiment, the individual operator standard deviations were examined to determine whether these discrepancies might be driven by one or two outliers in the distributions. In the principal HI–LO comparisons, the male σ's range from a low of 0.185 to a high of 1.476, with an average of 0.757, and the female σ's range from 0.407 to 1.689, with an average of 0.918. Four of the twenty males produce σ's in the HI–LO differences that exceed 1.000, while nine of the twenty female σ's exceed 1.000, thus indicating a clear tendency across the full operator pool toward



(9a)



(9b)

Fig. 9. Gender Comparisons in Local Pendulum Results.

TABLE 18
Composite Results of All Remote Pendulum Experiments

	All Operators	6 Male Oprs.	6 Female Oprs.
No. Run Sets	630	469	161
No. Series	126	93	32
HI Mean	–41682.06	–41650.67	–41773.52
Composite δ_c	1.682	1.651	1.774
Run Score σ	0.0666	0.0646	0.0735
Average δ_a	1.663	1.615	1.801
LO Mean	–41684.52	–41650.79	–41782.78
Composite δ_c	1.685	1.651	1.783
Run Score σ	0.0634	0.0598	0.0750
Average δ_a	1.670	1.625	1.802
BL Mean	–41685.70	–41653.58	–41779.27
Composite δ_c	1.686	1.654	1.779
Run Score σ	0.0824	0.0789	0.0898
Average δ_a	1.668	1.622	1.801
Comp. HI–LO (δ_{HL})	0.003	0.000	0.009
SD Diffs.	0.092	0.088	0.105
z-score	0.667	0.030	1.116
Prob.	0.252	0.488	0.132
Comp. HI–BL (δ_{HL})	0.004	0.003	0.006
SD Diffs.	0.106	0.102	0.116
z-score	0.860	0.616	0.625
Prob.	0.195	0.269	0.266
Comp. LO–BL (δ_{LB})	0.001	0.003	–0.004
SD Diffs.	0.104	0.099	0.117
z-score	(0.286)	(0.612)	–0.380
Prob.	(0.387)	(0.270)	0.352

1.000, thus indicating a clear tendency across the full operator pool toward larger variances in the female data.

9. Remote Pendulum Damping Experiments

Twelve operators, six males and six females, produced a smaller remote pendulum database of 126 series, or 630 sets, all following the five-set series format. Of these, 295 sets, or nearly half the total database, were generated by a single male operator. These results are summarized in Tables 18 and 19 and illustrated in Figure 10.

With only six operators in each group, interpretation of these results must be limited to the simple observations that the males have once again produced larger average deviations conforming to intention in the various comparisons, while the female composite databases display lower means and larger standard deviations.As in the local experiments, the reconstructed standard deviations produce significant F-ratios comparing male and female performance in all 3 intentions ($p_H = .02$; $p_L = 1\times10^{-4}$; $p_B = .02$), and the female composite baseline again has the largest individual σ.

TABLE 19
Individual Operator Results of All Remote Pendulum Experiments, by Gender

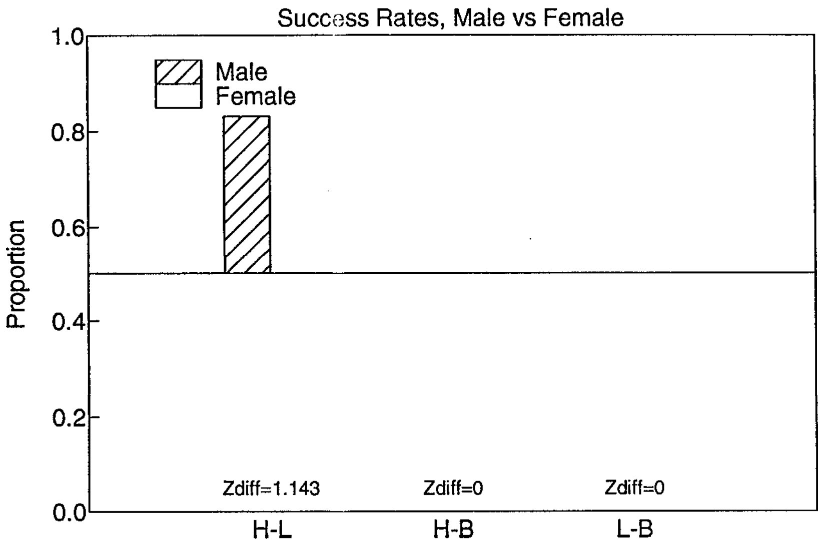
6 Male Operators			
	HI–LO	HI–BL	LO–BL
Avg # Run Sets	78.17	78.17	78.17
Avg Difference (δ)	0.010	0.007	–0.003
Avg *z*-Score	0.29	0.43	(0.18)
# Oprs *p*<0.50	5	3	3
Proportion	0.83	0.50	0.50
Prop. *z*-Score	1.63	0.00	0.00
# Oprs *p*<0.05	0 (0)	1 (0)	0 (0)
Proportion	0.00 (0.00)	0.17 (0.00)	0.00 (0.00)
Prop. *z*-Score	–0.64 (–0.64)	1.03 (–0.64)	–0.64 (–0.64)
6 Female Operators			
	HI–LO	HI–BL	LO–BL
Avg # Run Sets	26.83	26.83	26.83
Avg Difference (δ)	0.001	0.000	–0.000
Avg *z*-Score	0.13	0.10	–0.12
# Oprs *p*<0.50	3	3	3
Proportion	0.50	0.50	0.50
Prop. *z*-Score	0.00	0.00	0.00
# Oprs *p*<0.05	1 (0)	0 (0)	0 (0)
Proportion	0.17 (0.00)	0.00 (0.00)	0.00 (0.00)
Prop. *z*-Score	1.03 (–0.64)	–0.64 (–0.64)	–0.64 (–0.64)
Male/Female Diffs.			
z_{diff} p<0.50	1.14	0.00	0.00
z_{diff} p<0.05	–1.18 (0.00)	1.18 (0.00)	0.00 (0.00)

Combined Database

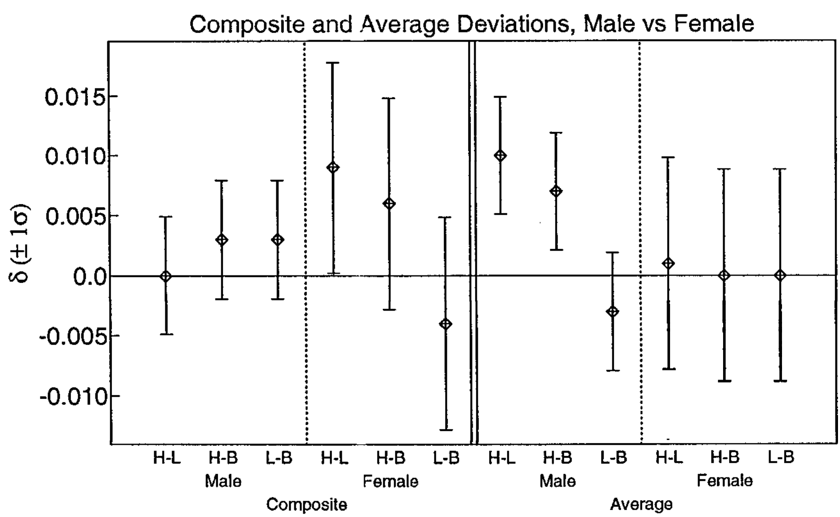
A. Proportional Comparisons

With the gender comparisons of all nine of these human/machine experiments calculated on commensurate differential measures, it becomes possible to combine their results to establish a more robust statistical assessment of the validity of some of the trends observed in the individual experiments. Of the various indicators that might be addressed in this combined database, comprising a total of 130 male and 140 female contributions, the most straightforward is a simple comparison of the overall proportions of operators in each group who achieve results consistent with their intentions. These are summarized in Table 20 and displayed in Figure 11.

These proportional comparisons confirm that across the full range of experiments there is a highly significant difference between the average male and female achievements, with the male operators outperforming the females in



(10a)



(10b)

Fig. 10. Gender Comparisons in Remote Pendulum Results.

TABLE 20
Combined Results of All PEAR Experiments, by Gender

	130 Male Operators		
	HI–LO	HI–BL	LO–BL
# Oprs p<0.50	82	74	69
Proportion	.63	.57	.53
Prop. z-Score	2.98*	1.58	0.70
# Oprs p<0.05	10 (3)	11 (1)	13 (4)
Proportion	0.08 (0.02)	0.08 (0.01)	0.10 (0.03)
Prop. z-Score	1.41 (–1.41)	1.81* (–2.21)†	2.62* (–1.01)
	140 Female Operators		
	HI–LO	HI–BL	LO–BL
# Oprs p<0.50	64	61	79
Proportion	0.46	0.44	0.56
Prop. z-Score	(–1.01)	(–1.52)	1.52
# Oprs p<0.05	9 (5)	10 (8)	8 (5)
Proportion	0.06 (0.04)	0.07 (0.06)	0.06 (0.04)
Prop. z-score	0.78 (–0.78)	1.16 (0.39)	0.39 (–0.78)
Male/Female Diffs.			
z_{diff} p<0.50	2.79*	2.13*	–0.49
Probability	0.003*	0.016*	0.311
z_{diff} p<0.05	0.33 (–0.33)	0.16 (–0.82)	0.66 (–0.16)
Probability	0.37 (0.37)	0.44 (0.21)	0.25 (0.44)

* and † — see Table Notes on p. 6.

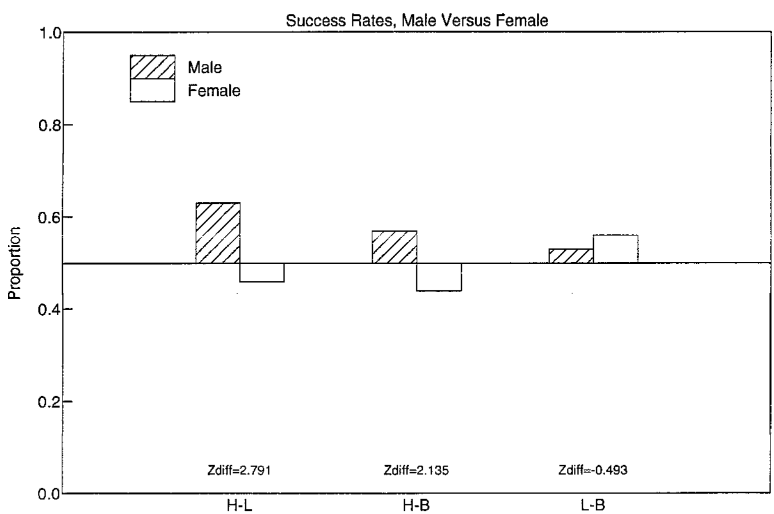


Fig. 11. Gender Comparisons for Combined Experimental Results.

the primary HI–LO comparison (z_{diff} = 2.79, p<.003). This effect is driven mainly by the substantial dissimilarities in the HI–BL proportions (z_{diff} = 2.13), with little difference in the LO–BL.

It is also apparent that this overall effect is the result of small but consistent effects generated by a majority of the operators, rather than by a few highly significant individual contributions. Of the total of 270 individual databases, only 19 (7%) exceed the p<.05 criterion in the HI–LO separations, ten males (8%) and nine females (6%), both little more than might be expected by chance. This figure is only slightly larger in the HI–BL and LO–BL comparisons, with a total of 21 significant achievements (8%) in each condition. Although 8% of the males exceed the p<.05 criterion in the HI–BL (z_M = 1.81) and 10% in the LO–BL (z_M = 2.62), none of the male/female differences are statistically significant. It is notable, however, that in all three comparisons the males show a deficit of results in the negative tails of the distributions, while the females produce a comparable number of extreme results in both tails, indicative of an overall shift in the intended directions in the male distributions in contrast to slightly larger scatters in the female distributions. Thus, despite the larger size and number of female contributions, and the fact that some of the strongest individual databases were generated by female operators, on average the females prove to be significantly less successful then the males in shifting the distribution means in accordance with their intentions.

B. Asymmetries

One of the more persistent gender-related patterns to emerge from the individual experiments is the apparent asymmetry of the intentional results relative to the baselines, particularly in the female performances. For example, of the three normalized average values for the high, low, and baseline intentions (or right, left, and baseline in the RMC experiments), indicated in the tables as δ_a, the females have their largest value in their baselines in six of the nine experiments, while the males produce their largest values in the high intentions in seven of the nine experiments. (Recall that these values are negative numbers in the pendulum experiments.) While these have no influence on the primary HI–LO comparisons, they do affect the comparative differences of the HI–BL and LO–BL calculations and contribute to the overall asymmetry of the composite databases. To assess this trend more quantitatively, an asymmetry parameter, A, defined as the proportion of operators in each group whose LO results are lower than their BL subtracted from the proportion whose HI results exceed their BL, has been calculated for both groups, and the male and female values of A compared.[4] The 130 male contributions, combined across

[4]Although the null hypothesis distribution of this A parameter is not intuitively obvious, it can be calculated and shown to be a function of N that rapidly approaches a normal distribution as N increases. With mean(A) = skew(A) = 0, standard deviation , and kurtosis(A) = 1.5/N, it follows that even moderate N's allow the direct calculation of a z-score by dividing A by the appropriate $\sigma(A)$.

all nine databases, yield $A_M = 0.04$ ($z_M = 0.56$) and the 140 female contributions $A_F = -.12$ ($z_F = -1.84$), resulting in a suggestive but non-significant difference between the two groups ($z_{M-F} = 1.62, p = 0.106$, two-tailed). However, if the null ATPseudo data are omitted, so that only those experiments displaying an overall anomalous effect are included in the calculation, the 110 male contributions then yield an $A_M = 0.07$ ($z_M = 0.93$), and the 120 female contributions an $A_F = -0.14$ ($z_F = -1.89$), with a marginally significant difference between the two groups ($z_{M-F} = 2.01, p = 0.044$). Thus, there is evidence for the existence of a stronger asymmetry in the female data that is statistically distinct from the male performance across the seven successful experiments. Since nearly 70% of the data presented in this survey were generated by female operators, this may well account, at least in part, for the persistent asymmetries observed across the various total experimental databases.

C. Residuals Analyses

The substantial variations in size of the individual and average databases in each group could conceivably distort these apparent gender-related differences. To address this possibility, residuals analyses were performed on each experimental database, under the null hypothesis that all operators produce the same statistical effect. Specifically, for every individual operator database the residuals from this common-effect hypothesis were calculated and sorted by gender.[5] Table 21 lists the resultant z_{M-F} and F-ratios, together with their associated probabilities, for each of the residuals comparisons of all nine experiments. (Probabilities are here calculated on a one-tailed basis since we are seeking confirmation of the hypothesis that the males produce larger residuals than the females.) These probabilities are then compounded using a standard meta-analytic formula, ($\chi^2 = -2\,\Sigma \log p_i$) [11], and evaluated via a χ^2 test with 18 d.f. These results are displayed in the bottom portion the table, along with those calculated for only the seven successful experiments (14 d.f.), for purposes of comparison.

With all experiments included, the combined results of these analyses indicate only a marginally significant difference between the two groups ($p = 0.047$), with the effect driven mainly by the strong HI–BL differences. The differences in the variances of the residuals distributions are indistinguishable from chance. If the null ATPseudo data are excluded, however, the gender differences in the remaining seven successful experiments are

[5]For each operator, the estimated effect size, ; the standard error of the estimate, ; and the operator residual, , where μ is the mean effect size for all operators , are calculated. If there is no difference in performance between the two groups, the R_i's should be z-distributed. To determine whether there is a difference between the two groups, we invoke to compare the average male residual (, where indicates the number of male operators and the sum is taken only over male contributions), with the average female residual .

TABLE 21
Male/Female Residuals Differences Across Nine Experiments

EXPERIMENT	HI–LO	HI–BL	LO–BL
REG (local)			
z_{M-F} (p_z)	1.550 (0.061)	0.888 (0.187)	–0.670 (0.749)
F-ratio (p_F)	0.539 (0.020)	0.813 (0.243)	1.488 (0.901)
REG (remote)			
z_{M-F} (p_z)	0.211 (0.416)	0.515 (0.303)	0.303 (0.381)
F-ratio (p_F)	0.783 (0.347)	0.648 (0.238)	0.815 (0.027)*
Pseudo (local)			
z_{M-F} (p_z)	0.539 (0.295)	1.786 (0.037)*	1.247 (0.106)
F-ratio (p_F)	1.251 (0.649)	5.473 (0.956)*	2.952 (0.872)
ATPseudo (local)			
z_{M-F} (p_z)	–0.691 (0.755)	–1.024 (0.847)	–0.331 (0.630)
F-ratio (p_F)	0.763 (0.301)	0.586 (0.158)	0.753 (0.293)
ATPseudo (remote)			
z_{M-F} (p_z)	–0.825 (0.795)	–0.538 (0.705)	0.365 (0.357)
F-ratio (p_F)	0.064 (0.061)	0.165 (0.148)	0.168 (0.151)
RMC (local)			
z_{M-F} (p_z)	0.066 (0.474)	1.200 (0.115)	1.126 (0.130)
F-ratio (p_F)	0.736 (0.276)	0.669 (0.217)	0.987 (0.498)
RMC (remote)			
z_{M-F} (p_z)	0.734 (0.232)	0.932 (0.176)	0.247 (0.402)
F-ratio (p_F)	0.928 (0.573)	1.377 (0.704)	0.010 (0.010)*
Pendulum (local)			
z_{M-F} (p_z)	2.832 (0.002)*	3.037 (0.001)*	0.550 (0.291)
F-ratio (p_F)	4.032 (0.998)*	1.589 (0.839)	1.315 (0.722)
Pendulum (remote)			
z_{M-F} (p_z)	0.196 (0.422)	0.404 (0.343)	0.454 (0.325)
F-ratio (p_F)	0.601 (0.295)	3.672 (0.910)	0.685 (0.344)
All Nine Experiments			
χ^2 z_{M-F} (p), 18 df	29.115 (0.047)*	36.757 (0.006)*	20.596 (0.300)
χ^2 F_{M-F}(p), 18 df	24.935 (0.127)	17.601 (0.482)	27.357 (0.073)
Excluding ATPseudo			
χ^2 z_{M-F} (p), 14 df	28.096 (.014)*	35.725 (0.008)*	17.134 (0.225)
χ^2 F_{M-F}(p), 14 df	16.957 (.258)	10.090 (0.756)	21.118 (0.099)

* — see Table Notes on p. 6.

considerably more pronounced (p = 0.014), with all of the male residuals larger than the corresponding female values in the primary HI–LO comparisons.

This meta-analytic strategy weights all the experiments equally, regardless of the number of participating operators. In an alternative approach, all the individual operator residuals may be pooled into a single distribution, thus weighting the experimental results by the number of contributing operators. This method was employed to construct the distributions displayed in Figure 12, where the curves are the Gaussian density functions having the same means and standard deviations as the respective male and female operator residuals. (Recall that the residuals are constructed to be normally distributed.) The 0 of the *x*-axis corresponds to the collective mean performance level, m, (see Note 5). Due to the individual normalization of each operator's

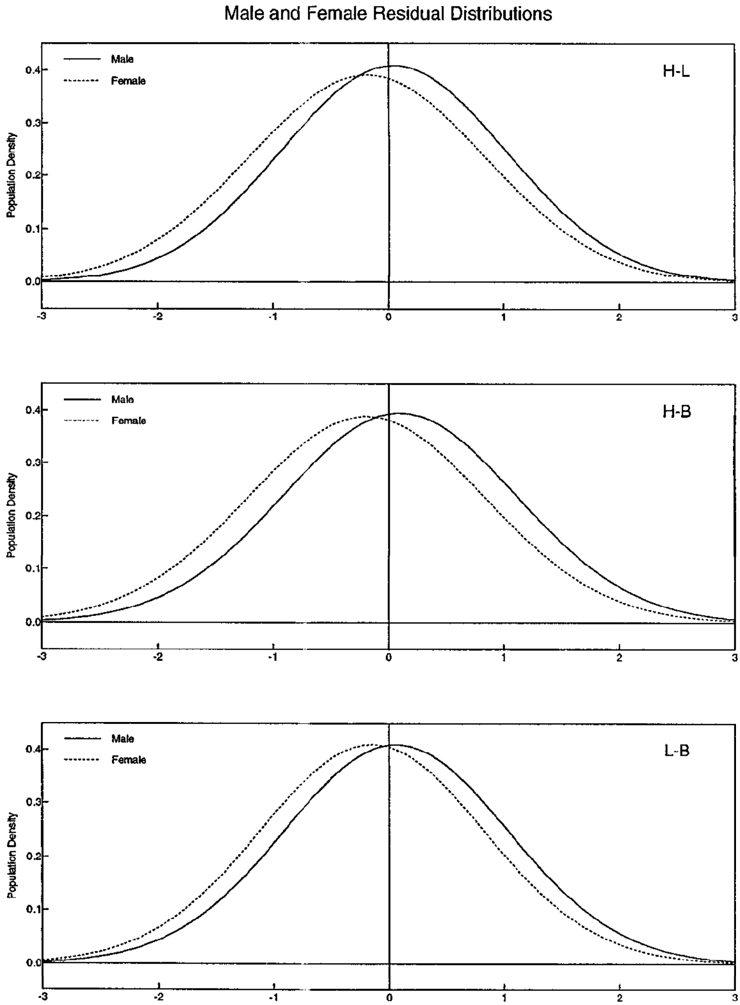


Fig. 12. Gender Comparisons of Pooled Residuals: All Experiments.

residual by that operator's standard error, the two distributions are not constrained to have their joint mean at 0, as would be the case for non-normalized residuals. Beyond providing a helpful graphic representation of the gender dissimilarities, the tabular results of the pooled residuals, noted in Table 22, confirm those of the proportional and meta-analytic computations in indicating significant differences in the primary HI–LO comparisons that are mainly attributable to the HI–BL performances. However, the LO–BL comparisons, which show no significant group differences in the proportional or meta-analytic calculations, also produce a marginally significant z_{diff} by this method, which may suggest dissimilarities at the individual operator level that cancel each other at the experimental level.

TABLE 22
Pooled Operator Residuals

All Nine Experiments			
	HI–LO	HI–BL	LO–BL
$z_{M\text{–}F}$ (prob.)	1.912 (.028)*	2.395 (.008)*	1.743 (.041)*
$F_{M\text{–}F}$ (prob.) (df= 129,139)	0.915 (.305)	0.966 (.421)	1.001 (.504)
Excluding ATPseudo Experiments			
	HI–LO	HI–BL	LO–BL
$z_{M\text{–}F}$ (prob.)	2.495 (.006)*	3.070 (.001)*	1.918 (.028)*
$F_{M\text{–}F}$ (prob.) (df= 109,119)	0.954 (.402)	1.065 (.632)	1.041 (.586)

* — see Table Notes on p. 6.

D. Remote vs. Local Comparisons

The residuals of the individual experiments listed in Table 21 also display apparent disparities between the yields of the local and remote experiments. To explore this more directly, the residuals of these two groups of experiments were calculated separately. Table 23 summarizes these results, derived by the meta-analytic method used to produce the values in Tables 21 and 22, both with and without the ATPseudo data.

The pooled residuals method produces comparable results for the five local and four remote experiments, as shown in Table 24, and were used to construct the distributions illustrated in Figures 13 and 14.

The results derived from both of these methods confirm the existence of significant gender differences in the local data, driven primarily by the differences in the HI–BL performances of the two groups, but show little evidence for consistent gender-related differences in the remote experiments. (Although the pooled residuals method fails to confirm the significance of the male/female disparities in the remote LO–BL variances indicated by the meta-analytic approach, these remain clearly evident in the graphs of Figure 14.) While it is possible that the small numbers of participating operators in the remote experiments may obscure some subtle gender differences (suggested by the consistency of larger male mean values in all the remote comparisons except the ATPseudo experiment), the lack of any statistical distinctions between the two groups in these remote databases is striking, given the highly significant differences in the local efforts, and may have important implications for comprehending the nature of the dissimilarities in gender performance.

TABLE 23
Comparison of Residuals of Remote *vs.* Local Experiments

	HI–LO	HI–BL	LO–BL
All Local Experiments			
χ^2_z (df)	22.254 (10)	28.052 (10)	12.539 (10)
p_z	0.014*	0.002*	0.251
χ^2_F(df)	13.687 (10)	10.020 (10)	4.987 (10)
p_F	0.188	0.439	0.892
All Remote Experiments			
χ^2_z (df)	6.860 (8)	8.705 (8)	8.057 (8)
p_z	0.552	0.368	0.428
χ^2_F (df)	11.255 (8)	7.581 (8)	22.371 (8)
p_F	0.188	0.475	0.004*
Local (ex. ATPseudo)			
χ^2_z (df)	21.694 (8)	27.720 (8)	11.614 (8)
p_z	0.006*	5×10^{-4}*	0.169
χ^2_F (df)	11.286 (8)	6.327 (8)	2.530 (8)
p_F	0.186	0.611	0.960†
Remote (ex. ATPseudo)			
χ^2_z (df)	6.402 (6)	8.005 (6)	5.999 (6)
p_z	0.380	0.238	0.423
χ^2_F (df)	5.671 (6)	3.763 (6)	18.589 (6)
p_F	0.461	0.709	0.005*

* and † — see Table Notes on p. 6.

TABLE 24
Pooled Operator Residuals

Five Local Experiments			
	HI–LO	HI–BL	LO–BL
z_{M-F} (prob.)	2.014 (0.022)*	2.284 (0.011)*	0.416 (0.339)
F_{M-F} (prob.) (df= 105,100)	0.923 (0.342)	1.002 (0.503)	1.242 (0.862)
Four Remote Experiments			
	HI–LO	HI–BL	LO–BL
z_{M-F} (prob.)	0.309 (.379)	0.818 (.207)	0.661 (.255)
F_{M-F} (prob.) (df= 22,38)	0.718 (.202)	0.810 (.300)	0.710 (.194)

* — see Table Notes on p. 6.

E. Standard Deviations

Finally, it is worth examining more closely the females' apparent tendency to produce larger standard deviations in a number of the experiments, both local and remote. While comparison of the trial standard deviations of the

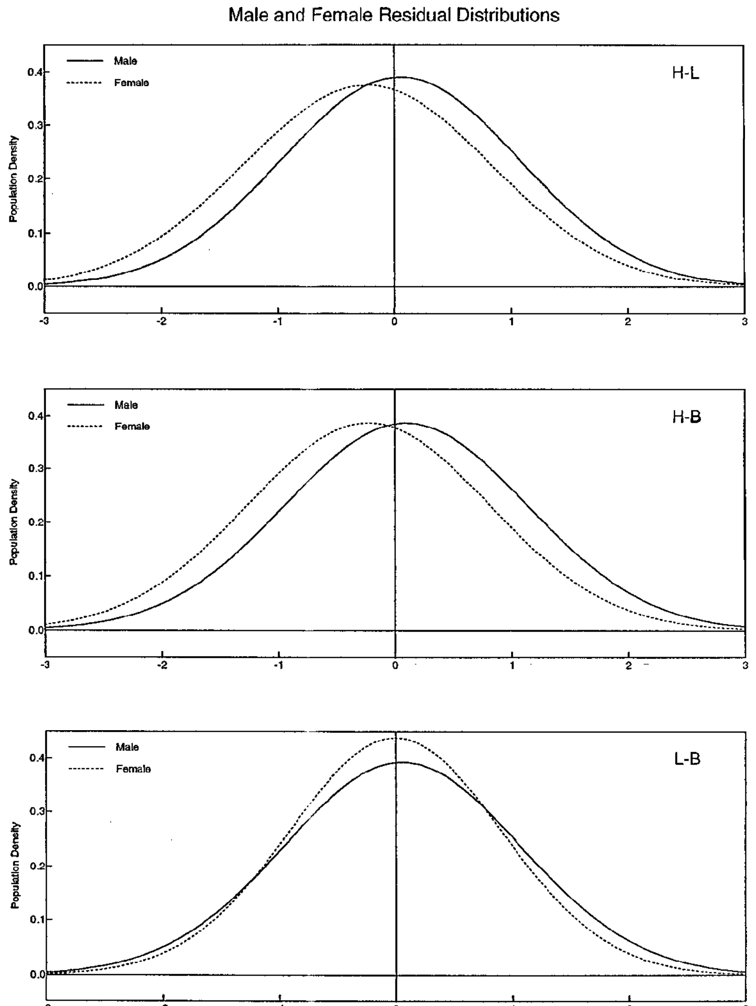


Fig. 13. Gender Comparisons of Pooled Residuals: Local Experiments.

REG-type experiments is quite straightforward, it should be recalled that the individual run standard deviations in the RMC and Pendulum experiments are confounded by spurious contributions from drifts of the means. This can be corrected by reconstructing the uncontaminated σ's of the single intentions from the standard deviations of the differential run comparisons provided in the various summary tables, following the procedure described in Note 3. Table 25 lists the F-ratios and associated one-tailed probabilities for the male vs female trial/run score variances of each intention for all nine experiments. These probabilities are compounded using the standard meta-analytic formula, with the χ^2 results displayed on the last line of the table.

These calculations leave little doubt about a significant gender-related dif-

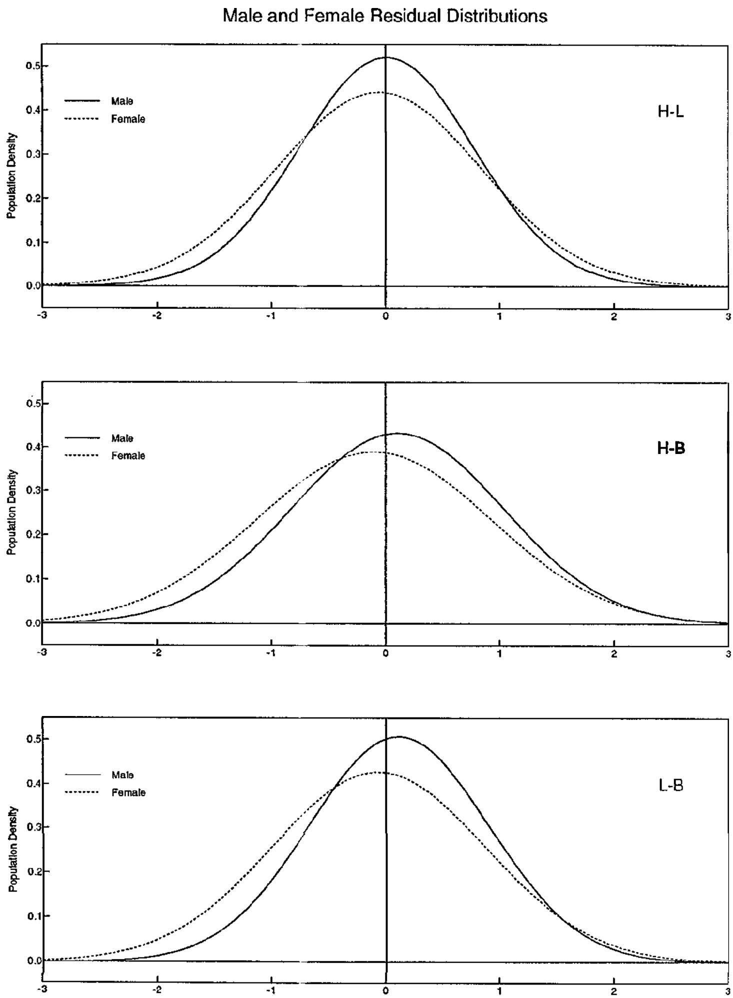


Fig. 14. Gender Comparisons of Pooled Residuals: Remote Experiments.

ference in the variances of these experimental data, even with the inclusion of the null ATPseudo experiments. (If these are omitted, the probabilities decrease by an order of magnitude in all three intentions.) Although the results of the high and baseline intentions are strongly influenced by the extreme female values in the Pendulum experiment and revert to chance when these are omitted, the differences in the low efforts remain highly significant ($p = .006$). Thus, it appears that although on average the females display relatively little success in shifting the means in the desired direction in their low efforts, they succeed in producing larger variances than the males in the output distributions in five of the seven successful experiments, cumulating to a statistically significant overall difference. This trend also manifests in the putatively null

TABLE 25
F-ratios of Male vs Female Trial/Run Score Variances

EXPERIMENT	HI	LO	BL
REG (local)			
F-ratio (df~330000,505000)	0.9952	0.9975	0.9955
Probability	0.064	0.214	0.080
REG (remote)			
F-ratio (df= 163999,293999)	1.0042	1.0011	0.9944
Probability	0.832	0.999	0.099
PseudoREG (local)			
F-ratio (df~12500,90000)	1.0163	1.0139	0.9991
Probability	0.886	0.848	0.475
ATPseudo (local)			
F-ratio (df= 76999,318999)	1.0042	1.0043	1.0031
Probability	0.770	0.776	0.708
ATPseudo (remote)			
F-ratio (df= 19999,65999)	0.9935	0.9921	0.9932
Probability	0.285	0 .244	0.276
RMC (local)			
F-ratio (df= 331,1008)	1.245	0.729	0.873
Probability	0.994	3×10^{-4}*	0.069
RMC (remote)			
F-ratio (df= 51,284)	0.798	0.617	1.045
Probability	0.166	0.019*	0.601
Pendulum (local)			
F-ratio (df= 305,608)	0.516	0.870	0.630
Probability	1×10^{-10}*	.084	3×10^{-6}*
Pendulum (remote)			
F-ratio (df= 468,161)	0.773	0.636	0.772
Probability	0.020*	1×10^{-4}	0.020*
χ^2 (d.f.)	66.620 (18)	54.292 (18)	54.087 (18)
Probability	2×10^{-7}*	2×10^{-5}*	1×10^{-5}*

* — see Table Notes on p. 6.

baseline efforts in six of the seven successful experiments. Given that their baselines means are also higher than those of the males in six of the seven successful experiments, these gender differences appear to reflect a fundamentally different mode of interaction with the various devices that may not be limited to the expression of simple conscious intention.

Summary and Discussion

Beyond providing statistical evidence for gender-related differences in performance in this genre of human/machine anomalies experiments, these analyses offer a number of specific indicators that may eventually be helpful in comprehending the basic source of the phenomena:

1. The female operators tend to generate larger databases than the males. Across the nine experiments, the 62 female databases constitute 69% of the data, compared to 31% from 73 male databases.

2. In contrast to the larger female composite deviations of the means in most of the databases, on an individual operator basis the males produce larger average deviations and corresponding *z*-scores.
3. Overall, the male operators are much more successful than the females in generating high-low separations consistent with their intentions.
4. The female databases display strong asymmetries in the two intentional directions of effort relative to their empirical baselines, possibly due in part to their tendency to displace the baselines from chance expectation.
5. Earlier evidence that the overall anomalous results are primarily attributable to small, statistically consistent shifts of the output distribution means produced by a majority of the operators, rather than to a few exceptional individual databases [1], is strongly reaffirmed in the male contributions, less so in the female.
6. The differences in male and female performance are much more distinct in the local experiments than in the remote.
7. Females tend to display larger variances than the males in their trial or run score distributions, an effect that manifests in their baselines as well as in their intentional efforts.
8. The overall null results of the ATPseudo databases apply equally to both genders, indicating that the gender-related patterns observed in the successful experiments are important components of the primary anomalies.

Collectively, these results indicate an underlying structure in the human/machine anomalies that is really related to some psychological, or possibly even physiological, characteristics of the human operators. Although the demonstrated gender-related patterns are only statistical indicators of group performance and hence limited in their capacity to predict individual achievement, they nonetheless raise a number of important questions regarding operator characteristics and experimental strategies that are well beyond the scope of this paper. These include the nature of the information processing dynamic that functions in such human/machine interactions, the psychological implications of "high," "low," and "baseline" intentions, and what is implied by the term "intention" itself. On this last point, another recent body of PEAR experiments, termed "FieldREG," has shown that anomalous human/machine effects can be produced in certain group environments in the absence of any conscious intentions, or even of conscious awareness, on the part of an operator [13]. These results, taken in conjunction with the co-operator outcomes that prompted the present study, suggest that "intention" may be only one contributing component of these phenomena, and that the ability of an individual to establish a resonant bond with another, or with a machine, may be a factor of comparable, or even greater, consequence.

These gender disparities may also hold important implications for the concept of a "baseline," or control condition, in any scientific study. The indications that many of the operators in these experiments, particularly the females, are producing distortions in the baselines, which are ostensibly non-intentional control conditions, raises questions about the generic reliability of such "controls." They also suggest that these anomalies may be associated with some deeper level of consciousness, one more closely identified with the brain's limbic functions than with its cognitive ones. In this vein, one might speculate that, at least for some people, the presence of a conscious intention may actually serve to inhibit the process that drives the basic phenomenon if it obstructs the subconscious resonance.

Finally, it is important to recognize that while this survey has focused on the distinction of biological gender, the variability in individual operator performances implies greater fundamental complexity of the phenomena. Any attempt to interpret these findings without taking into consideration such diverse variables as individual information-processing strategies, sociological expectations, technological sophistication, or personal belief systems, as well as a myriad of potential cultural and environmental factors that might influence performance in a task of this nature, will probably fall short of full understanding.

To conclude on the point with which we began, the strategy of the PEAR program to focus its efforts on the establishment of large databases has made it possible to detect a number of subtle structural sub-anomalies, such as these gender-related disparities, within the primary anomalous data distributions. Even so, the small signal-to-noise ratio of the primary effect makes it very difficult to address questions of structure or mechanism with the precision required to reach a fundamental understanding of the process. To do so would require a monumental effort to identify the most promising lines of inquiry, to design and implement an array of systematic studies capable of elucidating these elusive parameters, and to interpret their results incisively. But at the least, the results of this study offer some hope that such a program could be intellectually profitable.

Acknowledgements

The Princeton Engineering Anomalies Research program is indebted to the McDonnell Foundation, the Fetzer Institute, Mr. Laurance Rockefeller, Mr. Donald Webster, The Ohrstrom Foundation, the Institut für Grenzgebeite der Psychologie und Psychohygiene, The Lifebridge Foundation, and Mr. Richard Adams for their continued support of this research. We are also deeply appreciative of the enormous investments of time and energy by the many operators, male and female, who have generated these databases and contributed to our growing understanding through their suggestions and insights. Special thanks are extended to York Dobyns for his patient assistance in the various calculations and preparation of the numerous figures in this paper, and to Robert Jahn and Roger Nelson for their extensive analytical and editorial guidance.

References

[1] Dunne, B. J., Nelson, R. D., and Dobyns, Y. H. (1988). "Individual Operator Contributions in Large Data Base Anomalies Experiments," Technical Note PEAR 88002, Princeton Engineering Anomalies Research, Princeton University, School of Engineering/Applied Sciences.

[2] Dunne, B. J. (1991). "Co-Operator Experiments with an REG Device," Technical Note PEAR 91005, Princeton Engineering Anomalies Research, Princeton University, School of Engineering/ Applied Sciences, December 1991. [Published in modified form in *Cultivating Consciousness for Enhancing Human Potential, Wellness, and Healing,* K. R. Rao, ed. (Westport, CT and London: Praeger, 1993) p. 149–163.]

[3] Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H., and 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.

[4] Jahn, R. G., and Dunne, B. J. (1987). *Margins of Reality: The Role of Consciousness in the Physical World.* San Diego, New York, London: Harcourt Brace Jovanovich.

[5] Jahn, R. G., Dunne, B. J., and Nelson, R. D.(1987). Engineering anomalies research. *Journal of Scientific Exploration,* 1, 1, 21.

[6] Nelson, R. D. & Dobyns, Y. H. (1991). "Analysis of Variance of REG Experiments: Operator Intention, Secondary Parameters, Database Structure," Technical Note PEAR 91004, Princeton Engineering Anomalies Research, Princeton University, School of Engineering/ Applied Sciences.

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

[8] Dunne, B. J., Nelson, R. D,. and Jahn, R. G. (1988). Operator-related anomalies in a random mechanical cascade. *Journal of Scientific Exploration,* 2, 2, 155.

[9] Nelson, R. D., Bradish, G. J., Jahn, R. G., and Dunne, B. J. (1994). A linear pendulum experiment: Effects of operator intention on damping rate. *Journal of Scientific Exploration,* 8, 4, 471.

[10] Dunne, B. J. (1995). "Gender Differences in Human/Machine Anomalies," Technical Note PEAR 95005, Princeton Engineering Anomalies Research, Princeton University, School of Engineering and Applied Sciences.

[11] Rosenthal, R. (1984). *Meta-Analytic Procedures for Social Research.* Beverly Hills, CA: SAGE Publications, Inc.

[12] Dunne, B. J., Dobyns, Y. H., Jahn, R. G., and Nelson, R. D. (1994). Series position effects in random event generator experiments. With an Appendix by A. Thompson, Serial position effects in the psychological literature. *Journal of Scientific Exploration,* 8, 2, 197.

[13] Nelson, R. D., Bradish, G. J., Dobyns, Y. H., Dunne, B. J., and Jahn, R. G. (1996). FieldREG anomalies in group situations. *Journal of Scientific Exploration,* 10, 1, 111.