SRI International

Final Report--Objective I, Task 1
Covering the Period 31 July 1987 to 30 September 1988

December 1988

META-ANALYSIS OF FORCED-CHOICE PRECOGNITION EXPERIMENTS

By: EDWIN C. MAY
SRI International

CHARLES HONORTON
DIANE B. FERRARI
GEORGE HANSEN
Psychophysical Research Laboratories

Prepared for:

Peter J. McNelis, DSW
CONTRACTING OFFICER'S TECHNICAL REPRESENTATIVE

CONTRACT DAMD17-85-C-5130

SRI Project 1291

Approved by:

MURRAY J. BARON, Director
Geoscience and Engineering Center

SRI International®
333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

I INTRODUCTION

We have subcontracted to Psychophysical Research Laboratories (PRL) to conduct a meta-analysis of the forced-choice precognition literature. Mr. Honorton, the director, has met the requirements of the subcontract. Attached, is the deliverable from PRL.

Meta-analysis of Forced-choice Precognition Experiments

CHARLES HONORTON

DIANE BAILEY FERRARI

GEORGE HANSEN

Psychophysical Research Laboratories

301 College Road East, Princeton, N.J. 08540

Approved For Release 2000/08/08 : CIA-RDP96-00789R002200410001-2

TABLE OF CONTENTS

ABSTRACT

We report a meta-analysis of forced-choice precognition experiments published in English-language parapsychology journals between 1935 and 1987. These studies involve attempts by subjects to predict the identity or order of target stimuli selected randomly over intervals ranging from several hundred milliseconds to one year following the subject's responses. The database includes 309 studies reported by 62 senior authors. Nearly 2 million individual trials were contributed by more than 50,000 subjects. Study outcomes are assessed in terms of overall level of statistical significance and effect size.

We find a small, but consistent, and highly significant overall tendency for directional hitting ($z = 12.14$). Analysis based on investigators' predictions of conditions associated with hitting and missing yields a much stronger result ($z = 24.23$). Thirty percent of the studies (and 39% of the investigators) have directional outcomes that are significant at the 5% significance level. Assessment of the vulnerability of this database to selective reporting of positive results indicates that a ratio of 50 unreported studies averaging null results would be required for each reported study in order to reduce the overall significance of the observed outcomes to nonsignificance.

No systematic relationship exists between study outcomes and eight indices of research quality. Magnitude of effect has remained essentially constant over the survey period, while research quality has improved substantially.

Four moderating variables appear to covary significantly with study outcome:

- Studies using subjects selected on the basis of prior testing performance show significantly larger effects than studies involving unselected subjects.
- Subjects tested individually by an experimenter show significantly larger effects than those tested in groups.
- Studies in which subjects are given trial-by-trial or run-score feedback have significantly larger effects than those with limited or no subject feedback.
- Studies with brief intervals between subjects' responses and target generation show significantly stronger effects than studies involving longer intervals.

The combined impact of these moderating variables appears to be very strong. A nearly perfect replication rate is observed in the subset of studies using selected subjects, who are tested individually and receive trial-by-trial feedback.

OBJECTIVES

Precognition refers to the noninferential prediction of future events. Anecdotal claims of "future knowing" have occurred throughout human history, in virtually every culture and period. Today, such claims are generally believed to be based on factors such as delusion, irrationality, and superstitious thinking. The concept of precognition runs counter to accepted notions of causality and appears to conflict with current scientific theory. Nevertheless, over the past half-century, a substantial number of experiments have been reported by more than 60 investigators claiming empirical support for the hypothesis of precognition. Subjects in forced-choice experiments, according to many reports, have correctly predicted to a statistically significant degree the identity (or order) of target stimuli randomly selected at a later time.

We performed a meta-analysis of forced-choice precognition experiments published in the English-language research literature between 1935 and 1987. Five major questions were addressed through this meta-analysis:

- Is there overall evidence for accurate target identification (above chance hitting) in experimental precognition studies?
- Is there overall evidence that investigators can accurately predict tendencies toward hitting and missing?
- What is the magnitude of the overall (directional and predicted) precognition effect?
- Is the observed effect related to variations in methodological quality that could allow a more conventional explanation?
- Does precognition performance vary systematically with potential moderating variables, such as differences in subject populations, stimulus conditions, experimental setting, knowledge of results, and time interval between subject response and target generation?

DELINEATING THE DOMAIN

Source of Studies

Parapsychological research is still academically taboo and it is unlikely that there have been many dissertations and theses in this area that have escaped publication. Our retrieval of studies for this meta-analysis is therefore based on the published literature. The studies include all forced-choice precognition experiments appearing in the peer-reviewed English-language parapsychology journals: *Journal of Parapsychology, Journal (and Proceedings) of the Society for Psychical Research, Journal of the American Society for Psychical Research, European Journal of Parapsychology* (including the *Research Letter* of the Utrecht University Parapsychology Laboratory), and *Research in Parapsychology*.

Criteria for Inclusion

Our review is restricted to fixed length studies in which significance levels and effect sizes based on direct hitting can be calculated. Studies using outcome variables other than direct hitting, such as run-score variance and displacement effects, are included only if the report provides relevant information on direct hits (i.e., number of trials, hits, and probability of a hit). Finally, we exclude studies conducted by two investigators, S. G. Soal and Walter J. Levy, whose work has been unreliable.

Many published reports contain more than one experiment or experimental unit. Experiments involving multiple conditions are treated as separate study units.

Outcome Measures

Significance Levels: We calculated two significance estimates for each study. The *directional z-score* (z_{dir}) measures the subjects' success in scoring in the direction of their intention. The *predicted z-score* (z_{pred}) measures the investigator's success in predicting the relative strength or direction of the outcome through conditional comparisons, experimental manipulations, or correlations; above chance scoring (hitting) is assumed in single condition experiments unless psi-missing is explicitly predicted. Predicted z's have

positive signs if the study outcome supports the investigator's hypotheses and have negative signs if the outcome is opposite the investigator's hypotheses. The use of these two measures allows us to assess both overall accuracy (hitting) and lawfulness (predictability).

Effect Sizes: Most parapsychological experiments, particularly those in the older literature, have used the trial rather than the subject as the sampling unit. Thus, we must use a trial-based estimator of effect size. The effect size for each study is the *z*-score divided the square root of the number of trials in the study. As with significance levels, we have two effect sizes for each study. One reflects overall directional hitting (ES_{dir}) and the other is based on the investigators' predictions of hitting or missing (ES_{pred}).

General Characteristics of the Domain

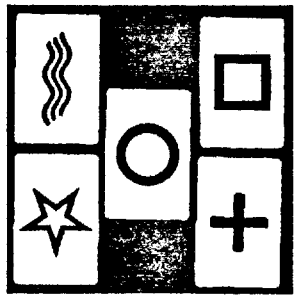
We located 309 studies in 113 separate publications. These studies were contributed by 62 different senior authors and were published over a 52-year period, between 1935 and 1987. Considering the half-century time-span over which the precognition experiments were conducted, it is not surprising that the studies are quite diverse.

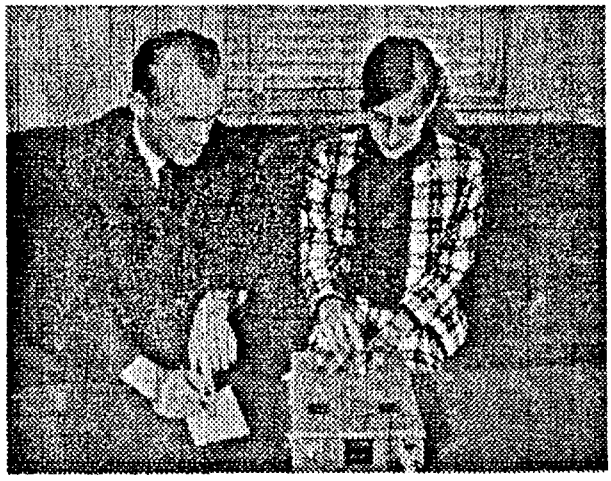
The data base comprises nearly 2 million individual trials and more than 50,000 subjects. Study sample sizes range from 25 to 297,060 trials (median = 1,194). The number of subjects ranges from 1 to 29,706 (median = 16). The studies employ a variety of methodologies, ranging from guessing Zener cards and other card symbols, to automated random number generator experiments (Figure 1). The domain encompasses diverse subject populations: the most frequently used population is students (used in approximately 40% of the studies); the least frequently used populations are the experimenters themselves and animals (each used in about 5% of the studies).

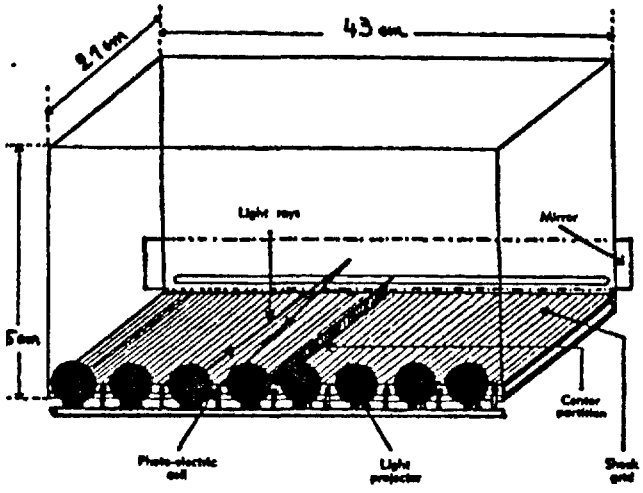
Though a few studies tested subjects through the mail, more typically subjects were tested in person, either individually or in groups. Target selection methods range from manual card-shuffling or dice-throwing to the use of random number tables or random number generators. The time-interval between the subjects' responses and target generation varies from less than one second to one year.

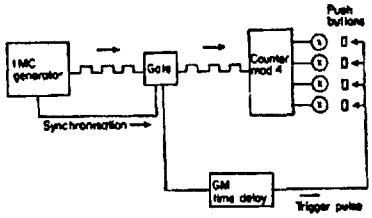
FIG. 1.: PRECOGNITION TESTING METHODS

TOP LEFT: Zener cards. TOP RIGHT: Subject tested with 4-choice random number generator. BOTTOM RIGHT: Block diagram for 4-choice RNG. BOTTOM LEFT: Apparatus used for small rodent anticipation of shock experiment.









OVERALL CUMULATION

Evidence for overall directional hitting and for successful prediction of hitting and missing tendencies is strong.[1] As shown in the top part of Table 1, the overall results are highly significant. The mean *predicted z* is twice as large as the mean *directional z*, indicating the advantage of making focused predictions (and the lawfulness implied by being able to do so). Thirty percent of the studies show overall significant hitting at the 5% level. Nearly 40% are significant on the basis of the investigators' predictions.

TABLE 1: Overall Precognition Significance Levels

	Z_{dir}	Z_{pred}
Mean	0.69	1.38
Standard Deviation	2.65	2.36
Combined (Stouffer) *z*	12.14	24.23
p_z	6×10^{-27}	4×10^{-52}
Filedrawer Estimate	16,529	66,687
Lower 95% Confidence Estimate	0.44	1.16
Lower 99% Confidence Estimate	0.34	1.07
Lower 99.9% Confidence Estimate	0.22	0.97

Lower-bound confidence estimates of the mean *z*-scores displayed in the bottom portion of Table 1 indicate that the mean *directional* and *predicted z-scores* are well above zero at the 95%, 99%, and 99.9% confidence levels.

Significance levels, not surprisingly, are related to sample size. The correlation (*r*) is 0.151 for the *directional z*'s (307 *df*, $p = .0044$), and for the

[1]The statistical analyses presented here were performed using SYSTAT (Wilkinson, 1988). When *t*-tests are reported on samples with unequal variances, they are calculated using the separate variances within groups for the error and degrees of freedom following Brownlee (1965). Unless otherwise specified, *p*-levels are one-tailed.

predicted z's, *r* is 0.242 (307 *df*, $p = 8.4 \times 10^{-6}$). Directionally significant studies have a mean sample size that is 37% larger than the mean for directionally nonsignificant studies. Using the *predicted z-score* criterion, significant studies have a mean sample size that is more than double that of the nonsignificant studies.

The effect size analysis is presented in Table 2. Both *directional* and *predicted* outcomes are significantly above zero, and again, the mean *predicted* effect size is twice as large as the *directional* mean *ES*.

TABLE 2: Overall Effect Sizes

	ES_{dir}	ES_{pred}
Mean	0.022	0.041
Standard Deviation	0.098	0.092
$t(308)$	4.01	7.88
p_t	4×10^{-5}	3×10^{-14}
Lower 95% Confidence Limit	0.012	0.032
Lower 99% Confidence Limit	0.008	0.029
Lower 99.9% Confidence Limit	0.005	0.025

Replication across Investigators

Virtually the same picture emerges when the cumulation is by investigator rather than study as the unit of analysis. The combined *z*'s are 12.71 for directional outcomes and 22.12 for predicted outcomes. Twenty-four of the 62 investigators (39%) have directional outcomes significant at the 5% level, and 39 investigators (63%) have significant predicted outcomes. The mean (investigator) *directional* effect size is 0.036 (*sd* = .091), and the mean *predicted ES* is 0.050 (*sd* = .087).

These results indicate a substantial level of cross-investigator replicability and directly contradict the claim of critics such as Akers (1988) that

successful parapsychological outcomes are achieved by only a few investigators.

The Filedrawer Problem

A well-known reporting bias exists throughout the behavioral sciences favoring publication of "significant" studies (e.g., Sterling, 1959). The extreme view of this "filedrawer problem," as Robert Rosenthal describes it, "is that the journals are filled with the 5% of the studies that show type I errors, while the filedrawers back at the lab are filled with the 95% of the studies that show nonsignificance. . ." (Rosenthal, 1984, p. 108). Recognizing the importance of this problem, the Parapsychological Association in 1975 adopted an official policy against selective reporting of positive results. Examination of the parapsychological literature shows that nonsignificant results are frequently published and in the precognition database, 60% to 70% of the studies have reported nonsignificant results. Nevertheless, 75% of the precognition studies were published before 1975, and we must ask to what extent selective publication bias could account for the cumulative effects we observe.[2]

The central section of Table 1 uses Rosenthal's (1984) filedrawer statistic to estimate the number of unreported studies with *z*-scores averaging zero that would be necessary to reduce the known database to nonsignificance. The filedrawer estimate suggests that over 50 unreported studies must exist for each reported study to reduce the cumulative hitting (directional) outcomes to a nonsignificant level. For the predicted outcomes, the filedrawer ratio is more than 200:1.

Another approach to the filedrawer problem is described by Robyn Dawes (Dawes, Landman and Williams, 1984; personal communication to Honorton, July 14, 1988). Dawes calculates the expected mean *z* and variance for various significance levels on the assumption that reported significant outcomes reflect nothing more than type I error. He then tests

[2]Analyses indicate no significant differences in the magnitude of reported study outcomes before and after 1975. The mean *directional* *z*-score for studies prior to 1975 is 0.719 (sd = 2.6) and for studies reported thereafter the mean is 0.605 (sd = 2.81) (t = 0.325, 307 df, p = .746, 2-tailed). For *predicted* *z*-scores, the comparable values are 1.43 (sd = 2.29) and 1.22 (sd = 2.60); t = 0.675, 307 df, p = .675, 2-tailed).

the difference between the observed and expected values. Applying this method to the precognition domain, it is extremely unlikely that the reported significance levels are just type I error. For the 5% significance level, for example, the mean observed and expected *directional z-scores* are 3.59 and 2.06, respectively. The observed mean is significantly larger than the expected value ($z = 4.10, p = .000021$). For the 0.5% significance level, the observed and expected means are 4.97 and 2.87 ($z = 7.0, p = 2.7 \times 10^{-12}$).

Based on these analyses, we conclude that the cumulative significance of the precognition studies cannot plausibly be attributed to selective reporting.

OUTLIER ELIMINATION

Although the overall significance levels and effect sizes for this database cannot reasonably be attributed to chance, inspection of the standard deviations in Tables 1 and 2 indicates that the study outcomes are extremely heterogeneous. Given the diversity of methods, subject populations, and other study features that characterize this research domain, this is not surprising.

The study outcomes are in fact extremely heterogeneous. Although a major objective of this meta-analysis is to account for the variability across studies by blocking on differences in study quality, procedural features, and sampling characteristics, the database clearly contains extreme outliers. The directional *z*-scores range from –5.06 to 19.6, a 25-sigma spread! The standardized index of kurtosis (g^2) is 9.86 ($p < 10^{-6}$), suggesting that the tails of the distribution are much too long for a normal distribution.

We have eliminated the extreme outliers by performing a "10-percent trim" on the study z-scores (Barnett and Lewis, 1978). This involves eliminating studies having *z*-scores in the upper and lower 10% of the distribution, and results in an adjusted sample of 248 studies. The directional *z*-scores for the adjusted sample range from -2.11 to 3.20 ($g2 = -1.1$). The revised significance levels and effect sizes are presented in Tables 3 and 4. Elimination of extreme outliers has reduced the combined significance levels by approximately one-half, but the outcomes remain highly significant. Twenty-five percent of the studies show overall significant hitting at the 5% level, and 28% are significant based on the investigators' predictions. Lower bound confidence estimates show that the *directional* and *predicted z*'s are above 0 at the 99.9% confidence level.

TABLE 3: Significance Levels for Adjusted Sample

	Z_{dir}	Z_{pred}
Mean	0.43	0.79
Standard Deviation	1.42	1.25
Combined (Stouffer) z	6.69	12.39
p_z	1.93×10^{-11}	1.36×10^{-27}
Lower 95% Confidence Estimate	0.28	0.66
Lower 99% Confidence Estimate	0.22	0.61
Lower 99.9% Confidence Estimate	0.15	0.55

Table 4 presents effect size estimates for the adjusted sample . Both the directional and predicted effect sizes remain significantly above zero.

TABLE 4: Effect Sizes for Adjusted Sample

	ES_{dir}	ES_{pred}
Mean	0.016	0.027
Standard Deviation	0.070	0.066
$t(247)$	3.60	6.44
p_t	1.92×10^{-4}	2.4×10^{-8}
Lower 95% Confidence Limit	0.009	0.020
Lower 99% Confidence Limit	0.006	0.017
Lower 99.9% Confidence Limit	0.002	0.014

Elimination of outliers reduces the total number of investigators from 62 to 57, but the results remain basically the same when the analyses are based on investigators rather than studies. The combined (Stouffer) z's are 7.37 for directional outcomes and 11.68 for predicted outcomes. Twenty one of the 57 investigators (36.8%) have directionally significant outcomes at the

5% level and 30 investigators (52.6%) have significant predicted outcomes. The mean (investigator) *directional* effect size is 0.023 (*sd* = .052), and the mean *predicted* effect size is 0.028 (*sd* = .047). Both results remain above 0 on lower-bound 99.9% confidence estimates.

Thus, elimination of the outliers does not substantially affect the conclusions drawn from our analysis of the database as a whole. There clearly is a nonchance effect. In the remainder of this report, we use the adjusted sample to examine covariations in magnitude of effect and a variety of methodological and other study features.

STUDY QUALITY

Study Quality Criteria

Since target stimuli in precognition experiments are selected only after the subjects' responses have been registered, precognition studies are usually not vulnerable to sensory leakage problems. Other potential threats to validity that must be, however, considered. For our analysis of study quality, statistical and methodological variables are defined and coded in terms of procedural descriptions (or their absence) in the research reports. One point is given (or withheld) for each of the following eight criteria:

Specification of Sample Size. Does the investigator preplan the number of trials to be included in the study or is the study vulnerable to the possibility of optional stopping? Credit is given to reports that explicitly specify the sample size. Studies involving group testing, in which it is not feasible to specify the sample size precisely, are also given credit. No credit is given to studies in which the sample size is either not preplanned or not addressed in the experimental report.

Preplanned Analysis. Is the method of statistical analysis, including the outcome (dependent variable) measure, preplanned? Credit is given to studies explicitly specifying the form of analysis and the outcome measure. No credit is given to those not explicitly stating the form of the analysis or those in which the analysis is clearly post hoc.

Randomization Method. Credit is given for use of random number tables, random number generators, and mechanical shufflers, but not for hand shuffling, die casting, or drawing lots.

Controls. Credit is given to studies reporting randomness control checks, such as random number generator (RNG) control series and empirical cross-check controls.

Recording. One point is allotted for automated recording of targets and responses and another for duplicate recording.

Checking. One point is allotted for automated checking of matches between target and response and another for duplicate checking of hits.

Study Quality Analysis

Each study received a quality weight between 0 and 8 (mean = 3.3, *sd* = 1.8). We find no relationship between study quality and effect size for either the directional (r_{246} = .029, p = .646, 2-tailed) or predicted (r_{246} = .006, p = .919, 2-tailed) effect sizes. Nor are any of the eight individual quality measures significantly related to effect size (Table 5).

TABLE 5: Point-biserial Correlations

Quality Measure	ES_{dir}	ES_{pred}
Sample size specified in advance	–.146	–.017
Preplanned analysis	–.042	–.002
Randomization	–.085	–.051
Controls	.036	.004
Automated recording	.139	–.016
Duplicate recording	.054	.074
Automated checking	.105	-.023
Duplicate checking	.015	.035

The mean effect sizes by quality level are displayed graphically in Figure 2 (*directional* outcomes) and Figure 3 (*predicted* outcomes).

FIGURE 2: Directional Outcomes in relation to Study Quality

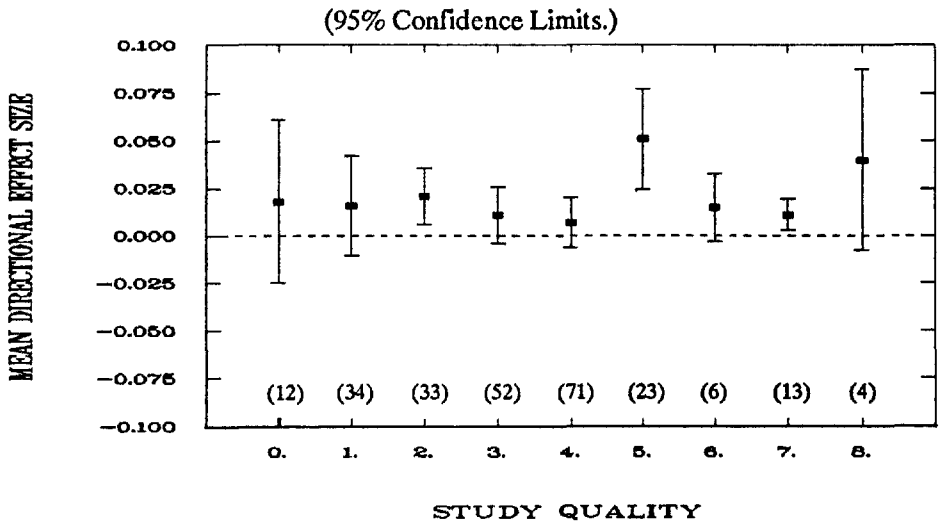
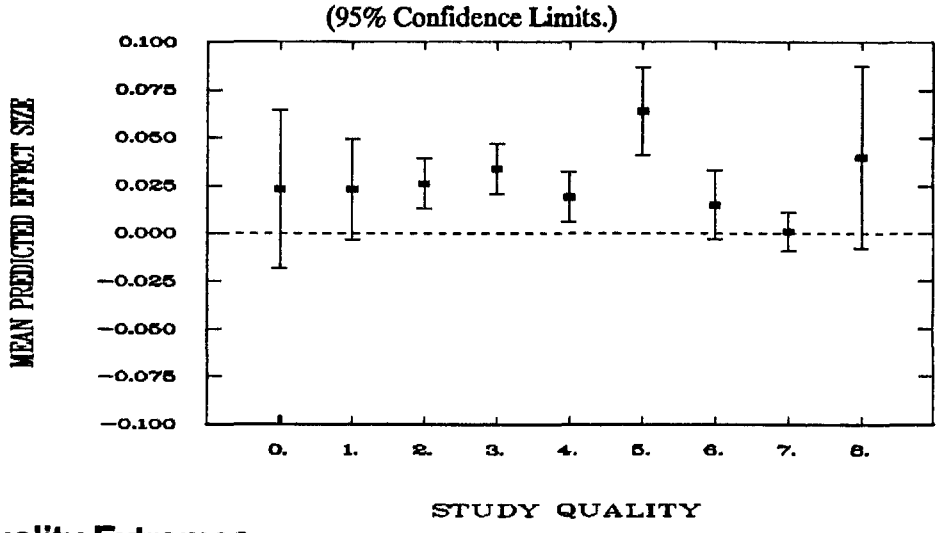


FIGURE 3: Predicted Outcomes in relation to Study Quality



Quality Extremes

Is there a tendency for extremely weak studies to show larger effects than exceptionally "good" studies? The grouped data presented in Figures 2 and 3 suggest that this is not the case and analysis on the extremes of the quality ratings indicates that the methodologically superior studies actually have somewhat *larger* mean effect sizes than studies with weaker methodology.

This analysis uses studies with quality ratings outside the interquartile range of the rating distribution (*median* = 3, Q_1 = 2, Q_3 = 4). There are 46 studies at each extreme ("low quality" = ratings of 0-1, "high quality" = ratings of 5-8). The high quality studies have *larger* effect sizes than the low quality studies in both the directional and predicted analyses. For the *directional* analysis, the effect size means are 0.034 (*sd* = 0.061) and 0.016 (*sd* = 0.091), for the high and low quality studies respectively (t = –1.09, 90 *df*, p = .278, 2-tailed). For the *predicted* analysis, the effect size means are 0.038 (*sd* = 0.059) and 0.023 (*sd* = 0.089), for the high and low quality studies respectively (t = –0.90, 90 *df*, p = .368, 2-tailed).

Quality Variation in Publication Sources

Study quality does vary significantly across the five publication sources. Although neither significance level nor effect size are significantly related to source of publication, the five journals do vary significantly in quality

(Kruskal-Wallis one-way ANOVA, chi-square = 11.78, 4 *df*, p = .019). This outcome is due to the substantially lower quality of studies appearing in the *Journal of the Society for Psychical Research.*

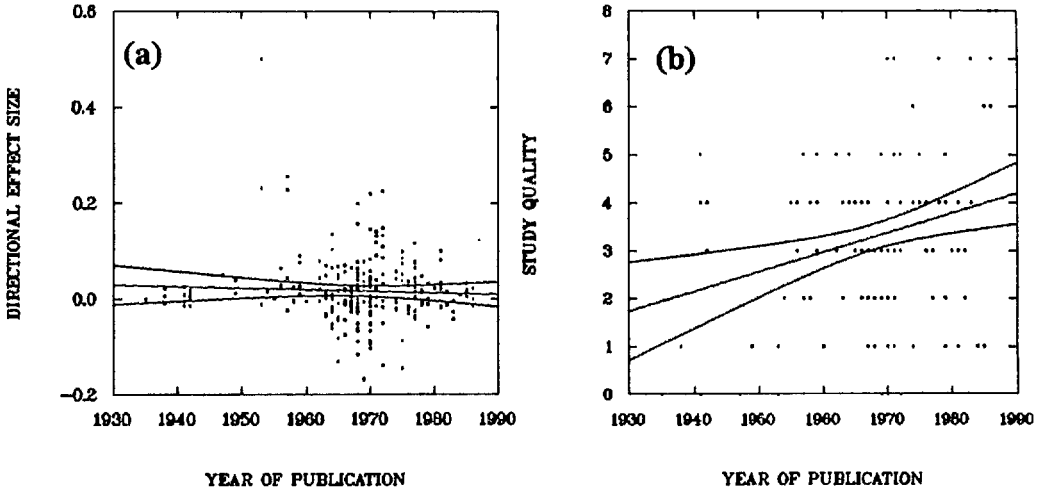
Study Quality in relation to Year of Publication

Precognition effect sizes have remained constant over a half-century of research, even though the methodological quality of the research has improved significantly during this period. The correlation between *directional* effect size and year of publication is –.050 (t_{246} = –0.79, p = .429). The result is nearly identical for the *predicted ES* (r_{246} = –.059, p = .358). Study quality and year of publication are, however, positively and significantly correlated (r_{246} = .239, p = 7.2 x 10^{-5}). See Figure 4.

Critics of parapsychology have long believed that evidence for parapsychological effects disappears as the methodological rigor increases. The precognition database does not support this belief.

FIGURE 4: (a) Directional Effect Sizes in relation to Year of Publication, (b) Study Quality in relation to Year of Publication

Least Squares Fit with 95% Confidence Limits



"REAL-TIME" ALTERNATIVES TO PRECOGNITON

Investigators have long been aware of the possibility that precognition effects could be modelled without assuming either time reversal or backward causality. For example, outcomes from studies with targets based on indeterminate random number generators (RNGs) could be due to a causal influence on the RNG—a psychokinetic (PK) effect—rather than information acquisition concerning its future state. In experiments with targets based on prepared tables of random numbers, the possibility exists that the experimenter or other randomizer may be the actual psi source, unconsciously using "real-time" ESP combined with PK to choose an entry point in the random number sequence that will significantly match the "subject's" responses. While this latter possibility may seem farfetched, it cannot be logically eliminated if one accepts the existing evidence for contemporaneous ESP and PK, and it has been argued that it is less farfetched than the alternative of "true" precognition.

Morris (1982) discusses models of experimental precognition based on "real-time" psi alternatives and methods for testing "true" precognition. In general terms these methods constrain the selection of the target sequence so as to eliminate nonprecognitive psi intervention. In the most common procedure, attributed to Mangan (1955), dice are thrown to generate a set of numbers which are mathematically manipulated to obtain an entry point in the random number table. This procedure is sufficiently complex "as to be apparently beyond the capacities of the human brain, thus ruling out PK because the 'PKer' would not know what to do even via ESP" (Morris, 1982, p. 329).

Two features of precognition study target determination procedures were coded to assess "real-time" psi alternatives to precognition:

- Method of determining random number table entry point,
- Use of Mangan's method.

Methods of eliminating "real-time" psi alternatives have not been employed in studies with random number generators and have only been

used in a small number of studies involving randomization by hand shuffling. These analyses are therefore restricted to studies using random number tables (N = 137).

Method of Determining RNT Entry Point

The reports describe six different methods of obtaining entry points in random number tables. If the study outcomes were due to subjects' precognitive functioning rather than to alternative psi modes on the part of the experimenter or the experimenter's assistants, there should be no difference in mean effect size across the various methods used to determine the entry point. Indeed, our analysis indicates that the study effect sizes do not vary systematically as a function of method of determining the entry point (Kruskal-Wallis one-way ANOVA by Ranks: chi-square = 8.29, 5 *df*, p = .141).

Use of Mangan's Method

We find no significant difference in effect size between studies using complex calculations of the type introduced by Mangan to fix the random number table entry point and those that do not use such calculations (t = 0.92, df = 77, p = .359, 2-tailed).

MODERATING VARIABLES

The stability of precognition study outcomes over a 50-year period is also bad news. It shows that investigators in this area have yet to develop sufficient understanding of the conditions underlying the occurrence (or detection) of these effects to reliably increase their magnitude. We have identified four variables that appear to covary systematically with magnitude of precognition performance:

- Selected versus unselected subjects
- Individual versus group testing
- Feedback level
- Time interval between subject response and target generation

We are interested only in factors associated with hitting; therefore, our analyses are based on the *directional* study outcomes only. The analyses use the raw study significance levels and effect sizes; this results in uniformly more conservative estimates of relationships with moderating variables than when the analyses are based on quality-weighted significance levels and effect sizes.

Selected versus Unselected Subjects

Our meta-analysis identifies eight subject populations:

- Unspecified subject populations
- Mixtures of several different populations
- Animals
- Students
- Children
- "Volunteers"
- Experimenter(s)
- Selected subjects

Effect size magnitude varies significantly across these eight subject populations (Kruskal-Wallis one-way ANOVA, chi-square = 15.71, 7 *df*, p = .028). Significance levels and effect sizes by subject population are displayed in Figures 5 and 6.

FIGURE 5: Significance Level by Subject Population

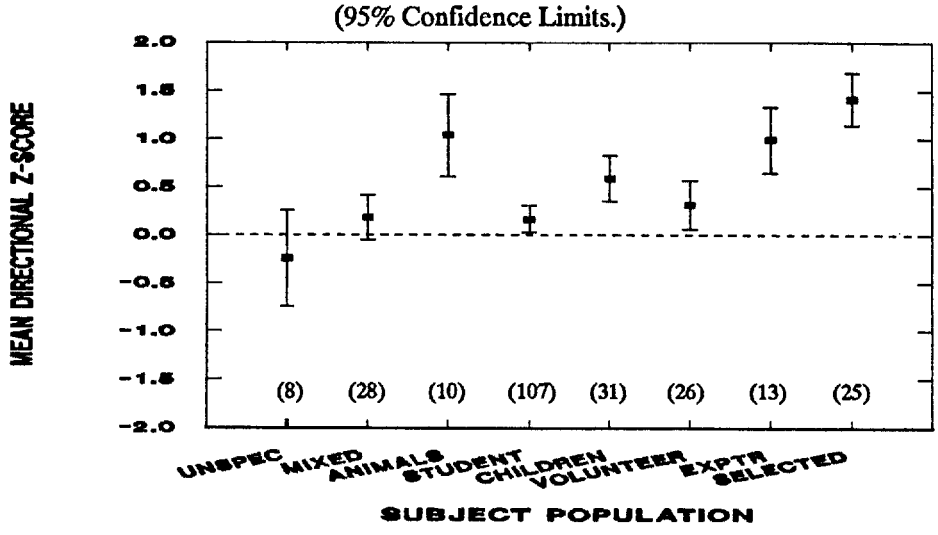
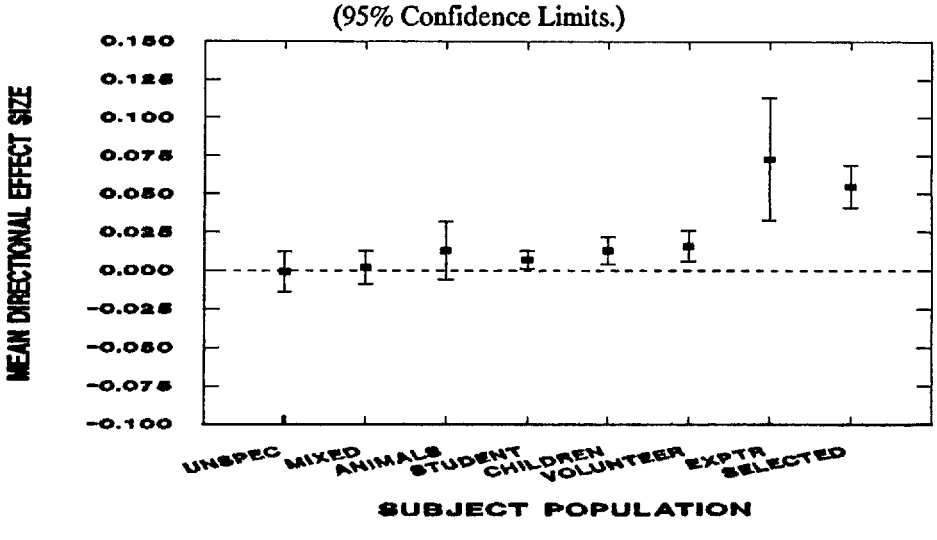


FIGURE 6: Effect Size by Subject Population



The difference across subject populations largely results from the superiority of studies with selected subjects: Studies using subjects selected on the basis of prior performance in experiments or pilot tests show larger effects than studies using unselected subjects. As shown in Table 6, 60% percent of the studies with selected subjects are significant at the 5% level. The mean z-score for these studies is 1.41 (sd = 1.36). The magnitude of effect size is significantly higher for selected subjects studies than for studies with unselected subjects. The t-test of the difference in mean effect size is equivalent to a point-biserial correlation of .186.

TABLE 6: Selected versus Unselected Subjects

Subjects	*N studies*	*Stouffer Z*	*Mean ES*$_{dir}$	*SD*	*% SIG.05*
Selected	25	7.05	0.055	0.072	60.0%
Unselected	223	4.70	0.012	0.068	21.0%

$t_{246} = 2.97, p = .0015$

Does this difference result from less stringent controls in studies with selected subjects? The answer appears to be "No." The average quality of studies with selected subjects is *higher* than studies using unselected subjects ($t_{27} = 2.05$, $p = .051$, 2-tailed). This result appears to reflect a general tendency toward increased rigor and more detailed reporting in studies with selected subjects.

Individual versus Group Testing

Subjects were tested in groups, individually, or through the mail. Studies in which subjects were tested individually by an experimenter have a significantly larger mean effect size than studies involving group testing (Table 7).

The t-test of the difference is equivalent to a point-biserial correlation of .234, favoring individual testing. Of the studies with subjects tested individually, 30.6% are significant at the 5% level.

The methodological quality of studies with subjects tested individually is significantly *higher* than that of studies involving group testing ($t_{143} = 3.5$, $p = .001$, 2-tailed). This result is consistent with the conjecture that group experiments are frequently conducted as "targets of opportunity," and may often be carried out hastily in an afternoon without the preparation and planning that goes into a study with individual subjects that may be conducted over a period of weeks or months.

TABLE 7: Individual versus Group Testing

Test Setting	*N studies*	*Stouffer Z*	*Mean ES*$_{dir}$	*SD*	*% SIG .05*
Individual	98	7.24	0.029	0.074	30.6%
Group	104	1.49	0.005	0.064	18.3%

$t_{200} = 2.40, p = .0085$

Thirty-five studies were conducted through the mail. In these studies, subjects completed the task at their leisure and mailed their responses to the investigator. These correspondence studies yield outcomes similar to those involving individual testing. The combined z-score is 3.01, with a mean effect size of 0.021 ($sd = .079$). Ten correspondence studies (28.6%) are significant at the 5% level.

Eleven studies are unclassifiable with regard to experimental setting.

Feedback

A significant positive relationship exists between the degree of feedback subjects receive about their performance and precognitive effect size (Table 8).

Subject feedback information is available for 95 studies. These studies fall into four feedback categories: No feedback, delayed feedback (usually notification by mail), run-score feedback, and trial-by-trial feedback. We gave these categories numerical values between 0 and 3. Precognition effect size correlates .258 with feedback level (103 $df, p = .004$). Of the 48 studies

involving trial-by-trial feedback, 21 (43.8%) are significant at the 5% level. None of the studies without subject feedback are significant.

TABLE 8: Feedback Received by Subjects

Feedback Level	*N studies*	*Stouffer z*	*Mean ES*$_{dir}$	*SD*$_{ES}$	*% SIG. 05*
No Feedback	15	0.00	-0.002	0.027	0.0%
Delayed	21	2.27	0.009	0.035	23.8%
Run-score	21	4.80	0.024	0.047	33.3%
Trial-by-trial	48	7.59	0.048	0.094	43.8%

While trial-by-trial feedback is associated with the largest effect sizes and significance levels, there is no evidence that subjects' performance improved over time.

Feedback level correlates positively though not significantly with research quality ($r_{103} = .134, p = .145$). Inadequate randomization is the most plausible source of potential artifacts in studies with trial-by-trial feedback. We therefore performed a separate analysis on the 48 studies in this group, blocking on the randomization and control quality measures. Studies with optimal randomization do not differ significantly in either mean significance level or mean effect size from those with suboptimal randomization. For significance levels, t is 0.74 with 46 df ($p = .465$, 2-tailed). For ES, t is 0.89 with 14 df ($p = .525$, 2-tailed). Similarly, studies reporting randomness control data do not differ significantly in either significance level or effect size from those not including randomness controls. For significance levels, t is 0.25 with 46 df ($p = .803$, 2-tailed). For ES, t is 1.19 with 46 df ($p = .241$, 2-tailed).

Time Interval

The interval between the subject's response and target selection ranges from less than one second to one year. Information about the time interval is available for 145 studies. This information, however, is often imprecise. Our analysis of the relationship between precognitive effect size and time interval is therefore limited to seven broad interval categories: milliseconds, seconds, minutes, hours, days, weeks, and months.

Although it is confounded with the feedback variable, there is a significant decline in precognition significance levels and effect size over increasing temporal distance. Using significance levels, r is –.270 with 143 df (p = .001, 2-tailed). Using effect size r is –.206 (p = .013, 2-tailed). The largest effects occur over the millisecond interval (N = 31 studies, Stouffer z = 6.12, mean ES = 0.046, sd = .072). The smallest effects occur over periods ranging from a week to a month (N = 17, Stouffer z = –.36, mean ES = –0.004, sd = .032).

Significance levels and effect sizes by precognitive interval are displayed in Figures 7 and 8. (The intervals are labelled numerically: 1 = msec., 2 = sec., 3 = min., 4 = hr., 5 = days, 6 = weeks, and 7 = months).

Curiously, this finding results entirely from studies using unselected subjects (r_{123} = –.238, p = .008, 2-tailed). Studies with selected subjects show a nonsignificant *positive* relationship betweem ES amd time interval (r_{18} = .081, p = .734, 2-tailed) and the difference between these two correlations is itself significant (z = 2.58, p = .01, 2-tailed). This suggests that the origin of the decline over time may be motivational rather than the result of some intrinsic physical boundary condition. The relationship between precognitive effect size and feedback also supports this conjecture. Nevertheless, any finding suggesting potential boundary conditions on the phenomenon should be vigorously pursued.

FIGURE 7: Significance Level by Precognitive Interval

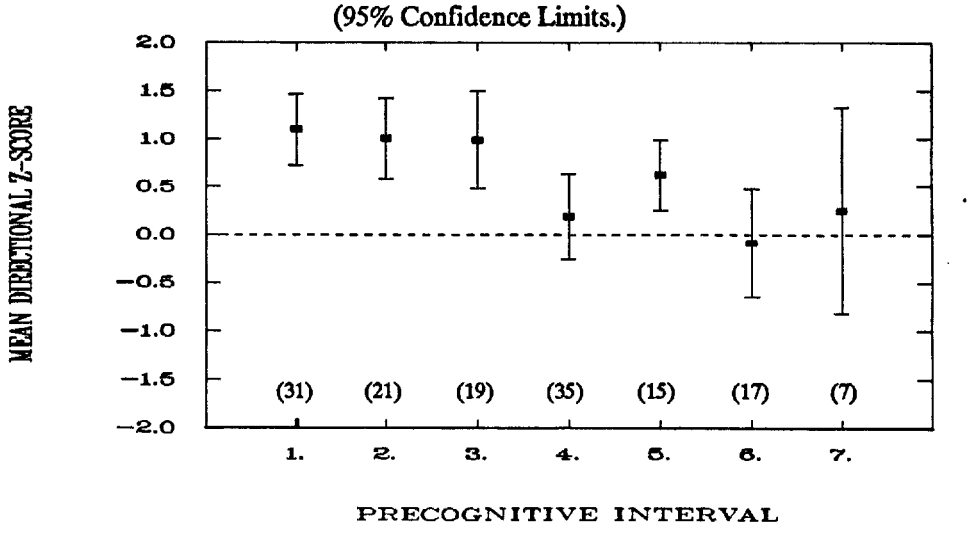
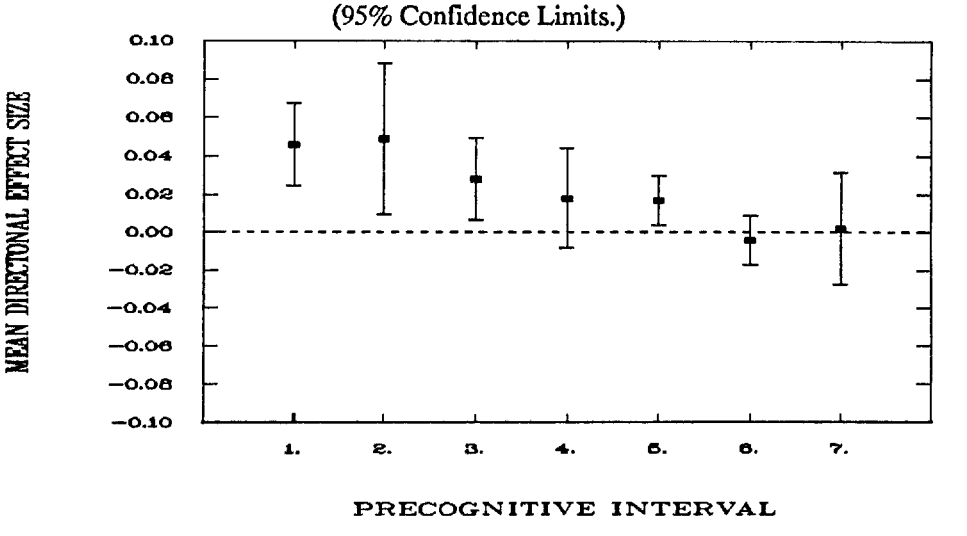


FIGURE 8: Effect Size by Precognitive Interval



Influence of Moderating Variables In Combination

The above analyses examine the impact of each moderating variable in isolation. In this final set of analyses, we explore their joint influence on precognition performance. For this purpose, we identify two subgroups of studies. One subgroup is characterized by the use of selected subjects tested individually with trial-by-trial feedback. We refer to this as the *Optimal* group (N = 8 studies). The second group is characterized by the use of unselected subjects tested in groups with no feedback. We refer to this as the *Suboptimal* group (N = 9 studies).

The Optimal studies are contributed by 4 independent investigators and the Suboptimal studies are contributed by 2 of same 4 investigators. All of the Optimal studies involve short precognitive time intervals (interval 1) while the Suboptimal studies involve longer intervals (intervals 5 and 6). All of the Optimal studies and 5 of the 9 Suboptimal studies use RNG methodology. The two groups do not differ significantly in average sample size. The mean study quality for the Optimal group is significantly higher than that of the Suboptimal studies (Optimal mean = 6.63, sd = 0.92; Suboptimal mean = 3.44, sd = 0.53; t = 8.63, 10 df, $p = 3.3 \times 10^{-6}$, 2-tailed).

The combined impact of the moderating variables appears to be quite strong: 7 of the 8 Optimal studies (87.5%) are independently significant at the 5% level, while none of the Suboptimal studies are statistically significant. All four investigators contributing studies to the Optimal group have significant outcomes. The mean z-score for the Optimal group is 2.17 (sd = 0.55) and for the Suboptimal group the mean z is -0.37 (sd = 1.05). The difference is highly significant (t = 6.13, 12 df, $p = 2 \times 10^{-5}$). The Optimal studies are also significantly less variable ($F_{(7,8)} = 3.67$, p = .046). In terms of effect sizes, the Optimal group is 9 times larger than the Suboptimal group (mean ES = 0.055, sd = 0.045 for the Optimal studies, and 0.006, sd = 0.033) for the Suboptimal studies; this difference is also significant (t = 2.60, 15 df, p = .01).

These findings suggest that future studies combining these moderators should yield especially promising outcomes.

DISCUSSION

Our meta-analysis of forced-choice precognition experiments confirms the existence of a small but statistically highly significant precognition effect. Most importantly, the effect appears to be replicable; significant confirmations are reported by two dozen investigators using a variety of methodological paradigms and subject populations.

Estimates of the "filedrawer" problem and consideration of parapsychological publication practices indicate that the precognition effect cannot be plausibly explained on the basis of selective publication bias. Analyses of precognitive effect sizes in relation to eight measures of research quality fail to support the hypothesis that the observed effect is driven to any appreciable extent by methodological artifacts; indeed, several of the analyses indicate that methodologically superior studies yield *stronger* effects than methodologically weaker studies.

Analyses of parapsychological alternatives to precognition, although limited to the subset of studies using random number tables, provide no support for the hypothesis that the effect results from the operation of contemporaneous ESP and PK at the time of randomization.

The most important outcome of the meta-analysis is the identification of several moderating variables that appear to covary systematically with precognition performance. The largest effects are observed in studies using subjects selected on the basis of prior test performance, who are tested individually, and receive trial-by-trial feedback. The outcomes of studies combining these factors contrast sharply with the null outcomes associated with the combination of group testing, unselected subjects, and no feedback of results. The identification of these moderating variables has important implications for our understanding of the phenomena and provides a clear direction for future research. The existence of moderating variables indicates that the precognition effect is not merely an unexplained departure from a theoretical chance baseline, but is rather an effect that covaries with factors known to influence more familiar aspects of human performance. It should now be possible to exploit these moderating factors to increase the magnitude and reliability of precognition effects in new studies.

While the overall precognition effect size is small, this does not imply that it has no practical consequences. It is, for example, of the same order of magnitude as effect sizes leading to the early termination of several major medical research studies. In 1981, the National Heart, Lung, and Blood Institute discontinued its study of propranolol because the results were so favorable to the propranolol treatment that it would be unethical to continue placebo treatment (Kolata, 1981); the effect size is 0.04. More recently, The Steering Committee of the Physicians' Health Study Research Group (1988), in a widely publicized report, terminated its study of the effects of aspirin in the prevention of heart attacks for the same reason. The aspirin group suffered 45% fewer heart heart attacks than a placebo control group; the associated effect size is 0.03.

The search for mechanisms underlying the phenomenon would be advanced considerably if it were possible to compare the magnitude of the precognition effect with the effect sizes in "real-time" ESP studies involving similar testing methods. Tart (1983) claims a robust and highly significant difference favoring "real-time" ESP in a small subset of forced-choice precognition and "real-time" ESP studies. However, his analysis is limited to 85 statistically significant studies (53 studies of "real-time" ESP and 32 precognition studies). Confirmation of this finding through comparative analysis of *all* retrievable "real-time" and precognition studies would have great value in efforts to model the phenomena and, also, for developing more effective research methods. Furthermore, although it is frequently claimed that ESP is independent of distance, we believe the evidence usually put forward in support of this claim is very weak and that a more satisfactory conclusion can only be reached through assessment of all of the evidence. For these reasons, we recommend that priority be given to a comprehensive meta-analysis of "real-time" ESP studies.

REFERENCES

Akers, C. (1987). Parapsychology is science, but its findings are inconclusive. *Behavioral and Brain Sciences*, **10**, 566-568.

Barnett, V., & Lewis, T. (1978). *Outliers in statistical data.* New York: John Wiley & Sons, Inc.

Brownlee, K. A. (1965). Statistical theory and methodology in science and engineering. New York: John Wiley & Sons, Inc.

Dawes, R., Landman, J., & Williams, M. (1984). Reply to Kurosawa. *American Psychologist*, **38**, 74-75.

Kolata, G. B. (1981). Drug found to help heart attack survivors. *Science*, **214**, 774-775.

Morris, R. L. (1982). Assessing experimental support for true precognition. *Journal of Parapsychology*, **46**, 321-336.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance--or vice versa. *Journal of the American Statistical Association*, **54**, 30-34.

Tart, C. T. (1983). Information acquisition rates in forced-choice ESP experiments: precognition does not work as well as present-time ESP. *Journal of the American Society for Psychical Research*, **77**, 293-310.

The Steering Committee of the Physicians' Health Study Research Group. (1988). Preliminary report: Findings from the aspirin component of the ongoing Physicians' Health Study. *New England Journal of Medicine*, **318**, 262-264.

Wilkinson, L. (1988). *SYSTAT: The system for statistics.* Evanston, IL: SYSTAT, inc.

CHRONOLOGICAL LISTING OF META-ANALYSIS STUDY REFERENCES

1 CARINGTON, W. (1935). Preliminary experiments in precognitive guessing. *Journal of the Society for Psychical Research*, **29**, 86-104.

2-5 RHINE, J. B. (1938). Experiments bearing on the precognition hypothesis: I. Pre-shuffling card calling. *Journal of Parapsychology*, **2**, 38-54.

6-11 RHINE, J. B., SMITH, B. M., & WOODRUFF, J. L. (1938). Experiments bearing on the precognition hypothesis: II. The role of ESP in the shuffling of cards. *Journal of Parapsychology*, **2**, 119-131.

12 HUMPHREY, B. M., & PRATT, J. G. (1941). A comparison of five ESP test procedures. *Journal of Parapsychology*, **5**, 267-293.

13-16 RHINE, J. B. (1941). Experiments bearing upon the precognition hypothesis: III. Mechanically selected cards. *Journal of Parapsychology*, **5**, 1-57.

17 STUART, C. E. (1941). An analysis to determine a test predictive of extra-chance scoring in card-calling tests. *Journal of Parapsychology*, **5**, 99-137.

18-21 HUMPHREY, B. M., & RHINE, J. B. (1942). A confirmatory study of salience in precognition tests. *Journal of Parapsychology*, **6**, 190-219.

22-25 RHINE, J. B. (1942). Evidence of precognition in the covariation of salience ratios. *Journal of Parapsychology*, **6**, 111-143.

26 NICOL, J. F., & CARINGTON, W. (1947). Some experiments in willed die-throwing. *Proceedings of the Society for Psychical Research*, **48**, 164-175.

27-28 THOULESS, R. H. (1949). A comparative study of performance in three psi tasks. *Journal of Parapsychology*, **13**, 263-273.

29-32 BASTIN, E. W., & GREEN, J. M. (1953). Some experiments in precognition. *Journal of Parapsychology*, **17**, 137-143.

33 MCMAHAN, E. A., & BATES, E. K. (1954). Report of further marchesi experiments. *Journal of Parapsychology*, **18**, 82-92.

34 MANGAN, G. L. (1955). Evidence of displacement in a precognition test. *Journal of Parapsychology*, **19**, 35-44.

35-36 OSIS, K. (1955). Precognition over time intervals of one to thirty-three days. *Journal of Parapsychology*, **19**, 82-91.

37 NIELSEN, W. (1956). An exploratory precognition experiment. *Journal of Parapsychology*, **20**, 33-39.

38-39 NIELSEN, W. (1956). Mental states associated with success in precognition. *Journal of Parapsychology*, **20**, 96-109.

40-41 FAHLER, J. (1957). ESP card tests with and without hypnosis. *Journal of Parapsychology*, **21**, 179-185.

42-44 MANGAN, G. L. (1957). An ESP experiment with dual-aspect targets involving one trial a day. *Journal of Parapsychology*, **21**, 273-283.

45-46 ANDERSON, M., & WHITE, R. (1958). A survey of work on ESP and teacher-pupil attitudes. *Journal of Parapsychology*, **22**, 246-268.

47 NASH, C. B. (1958). Correlation between ESP and religious value. *Journal of Parapsychology*, **22**, 204-209.

48-49 ANDERSON, M. (1959). A precognition experiment comparing time intervals of a few days and one year. *Journal of Parapsychology*, **23**, 81-89.

50-53 ANDERSON, M., & GREGORY, E. (1959). A two-year program of tests for clairvoyance and precognition with a class of public school pupils. *Journal of Parapsychology*, **23**, 149-177.

54 NASH, C. B. (1960). Can precognition occur diametrically? *Journal of Parapsychology*, **24**, 26-32.

55 FREEMAN, J. A. (1962). An experiment in precognition. *Journal of Parapsychology*, **26**, 123-130.

56 RHINE, J. B. (1962). The precognition of computer numbers in a public test. *Journal of Parapsychology*, **26**, 244-251.

57 RYZL, M. (1962). Training the psi faculty by hypnosis. *Journal of the Society for Psychical Research*, **41**, 234-252.

58-60 SANDERS, M. S. (1962). A comparison of verbal and written responses in a precognition experiment. *Journal of Parapsychology*, **26**, 23-34.

61-64 FREEMAN, J. (1963). Boy-girl differences in a group precognition test. *Journal of Parapsychology*, **27**, 175-181.

65-68 RAO, K. R. (1963). Studies in the preferential effect II. A language ESP test involving precognition and "intervention". *Journal of Parapsychology*, **27**, 147-160.

69-70 FREEMAN, J. (1964). A precognition test with a high-school science club. *Journal of Parapsychology*, **28**, 214-221.

71-73 FREEMAN, J., & NIELSEN, W. (1964). Precognition score deviations as related to anxiety levels. *Journal of Parapsychology*, **28**, 239-249.

74-82 SCHMEIDLER, G. (1964). An experiment on precognitive clairvoyance. Part I. The main results. *Journal of Parapsychology*, **28**, 1-14.

83-90 FREEMAN, J. A. (1965). Differential response of the sexes to contrasting arrangements of ESP target material. *Journal of Parapsychology*, **29**, 251-258.

91-92 OSIS, K., & FAHLER, J. (1965). Space and time variables in ESP. *Journal of the American Society for Psychical Research*, **59**, 130-145.

93-94 FAHLER, J., & OSIS, K. (1966). Checking for awareness of hits in a precognition experiment with hypnotized subjects. *Journal of the American Society for Psychical Research*, **60**, 340-346.

95-102 FREEMAN, J. A. (1966). Sex differences and target arrangement: High-school booklet tests of precognition. *Journal of Parapsychology*, **30**, 227-235.

103 ROGERS, D. P. (1966). Negative and positive affect and ESP run-score variance. *Journal of Parapsychology*, **30**, 151-159.

104 ROGERS, D. P., & CARPENTER, J. C. (1966).. The decline of variance of ESP scores within a testing session. *Journal of Parapsychology*, **30**, 141-150.

105 BRIER, B. (1967). A correspondence ESP experiment with high-I.Q. subjects. *Journal of Parapsychology*, **31**, 143-148.

106-107 BUZBY, D. E. (1967). Precognition and a test of sensory perception. *Journal of Parapsychology*, **31**, 135-142.

108-109 BUZBY, D. E. (1967). Subject attitude and score variance in ESP tests. *Journal of Parapsychology*, **31**, 43-50.

110-113 FREEMAN, J. A. (1967). Sex differences, target arrangement, and primary mental abilities. *Journal of Parapsychology*, **31**, 271-279.

114-118 HONORTON, C. (1967). Creativity and precognition scoring level. *Journal of Parapsychology*, **31**, 29-42.

119-120 CARPENTER, J. C. (1968). Two related studies on mood and precognition run-score variance. *Journal of Parapsychology*, **32**, 75-89.

121 DUVAL, P., & MONTREDON, E. (1968). ESP experiments with mice. *Journal of Parapsychology*, **32**, 153-166.

122-129 FEATHER, S. R., & BRIER, R. (1968). The possible effect of the checker in precognition tests. *Journal of Parapsychology*, **32**, 167-175.

130-137 FREEMAN, J. A. (1968). Sex differences and primary mental abilities in a group precognition test. *Journal of Parapsychology*, **32**, 176-182.

138-139 NASH, C. S., & NASH, C. B. (1968). Effect of target selection, field dependence, and body concept on ESP performance. *Journal of Parapsychology*, **32**, 248-257.

140-143 RHINE, L. E. (1968). Note on an informal group test of ESP. *Journal of Parapsychology*, **32**, 47-53.

144-146 RYZL, M. (1968). Precognition scoring and attitude toward ESP. *Journal of Parapsychology*, **32**, 1-8.

147-148 RYZL, M. (1968). Precognition scoring and attitude. *Journal of Parapsychology*, **32**, 183-189.

149-150 CARPENTER, J. C. (1969). Further study on a mood adjective check list and ESP run-score variance. *Journal of Parapsychology*, **33**, 48-56.

151 DUVAL, P., & MONTREDON, E. (1969). Precognition in mice: A confirmation. *Journal of Parapsychology*, **33**, 71-72.

152-155 FREEMAN, J. A. (1969). The psi-differential effect in a precognition test. *Journal of Parapsychology*, **33**, 206-212.

156 FREEMAN, J. A. (1969). A precognition experiment with science teachers. *Journal of Parapsychology*, **33**, 307-310.

157-158 JOHNSON, M. (1969). Attitude and target differences in a group precognition test. *Journal of Parapsychology*, **33**, 324-325.

159 MONTREDON, E., & ROBINSON, A. (1969). Further precognition work with mice. *Journal of Parapsychology*, **33** , 162-163.

160-162 SCHMIDT, H. (1969). Precognition of a quantum process. *Journal of Parapsychology*, **33**, 99-108.

163 BENDER, H. (1970). Differential scoring of an outstanding subject on GESP and clairvoyance. *Journal of Parapsychology*, **34**, 272-273.

164-165 FREEMAN, J. (1970). Shift in scoring direction with junior-high-school students: A summary. *Journal of Parapsychology*, **34**, 275.

166-169 FREEMAN, J. A. (1970). Ten-page booklet tests with elementary-school children. *Journal of Parapsychology*, **34**, 192-196

170-171 FREEMAN, J. A. (1970). Mood, personality, and attitude in precognition tests. *Journal of Parapsychology*, 34,. 322

172-175 FREEMAN, J. A. (1970). Sex differences in ESP response as shown by the Freeman picture-figure test. *Journal of Parapsychology*, **34**, 37-46.

176-179 HARALDSSON, E. (1970). Subject selection in a machine precognition test. *Journal of Parapsychology*, **34**, 182-191.

180 HARALDSSON, E. (1970). Precognition of a quantum process: A modified replication. *Journal of Parapsychology*, **34**, 329-330.

181-200 NIELSEN, W. (1970). Relationships between precognition scoring level and mood. *Journal of Parapsychology*, **34**, 93-116.

201-202 SCHMIDT, H. (1970). Precognition test with a high-school group. *Journal of Parapsychology*, **34**, 70.

203-204 BELOFF, J., & BATE, D. (1971). An attempt to replicate the Schmidt findings. *Journal of the Society for Psychical Research*, **46**, 21-31.

205-206 HONORTON, C. (1971). Automated forced-choice precognition tests with a "sensitive". *Journal of the American Society for Psychical Research*, **65**, 476-481.

207 MITCHELL, E. D. (1971). An ESP test from Apollo 14. *Journal of Parapsychology*, **35**, 89-107.

208-209 SCHMIDT, H., & PANTAS, L. (1971). Psi tests with psychologically equivalent conditions and internally different machines. *Journal of Parapsychology*, **35**, 326-327.

210 STANFORD, R. G. (1971). Extrasensory effects upon "memory". *Journal of the American Society for Psychical Research*, **64**, 161-186.

211-212 STEILBERG, B. J. (1971). Investigation of the paranormal gifts of the Dutch sensitive Lida T. *Journal of Parapsychology*, **35**, 219-225.

213 THOULESS, R. H. (1971). Experiments on psi self-training with Dr. Schmidt's pre-cognitive apparatus. *Journal of the Society for Psychical Research*, **46**, 15-21.

214-215 CRAIG, J. G. (1972). The effect of contingency on precognition in the rat. *Research in Parapsychology*, 154-156.

216-217 FREEMAN, J. A. (1972). The psi quiz: A new ESP test. *Research in Parapsychology*, 132-134.

218-220 HONORTON, C. (1972). Reported frequency of dream recall and ESP. *Journal of the American Society for Psychical Research*, **66**, 369-374.

221 JOHNSON, M., & NORDBECK, B. (1972). Variation in the scoring behavior of a "psychic" subject. *Journal of Parapsychology*, **36**, 122-132.

222 KELLY, E. F., & KANTHAMANI, B. K. (1972). A subject's efforts toward voluntary control. *Journal of Parapsychology*, **36**, 185-197.

223-226 SCHMIDT, H., & PANTAS, L. (1972). Psi tests with internally different machines. *Journal of Parapsychology*, **36**, 222-232.

227 ARTLEY, B. (1974). Confirmation of the small-rodent precognition work. *Journal of Parapsychology*, **38**, 238-239.

228-230 HARALDSSON, E. (1974).. Reported dream recall, precognitive dreams, and ESP. *Research in Parapsychology*, 47-48.

231 HARRIS, S., & TERRY, J. (1974) Precognition in a water-deprived Wistar rat. *Journal of Parapsychology*, **38**, 239.

232 RANDALL, J. L. (1974). An extended series of ESP and PK tests with three English schoolboys. *Journal of the Society for Psychical Research*, **47**, 485-494.

233 TERRY, J. C., & HARRIS, S. A. (1974). Precognition in water-deprived rats. *Research in Parapsychology*, 81.

234-235 EYSENCK, H. J. (1975) Precognition in rats. *Journal of Parapsychology*, **39**, 222-227.

236-237 HONORTON, C., RAMSEY, M., & CABIBBO, C. (1975). Experimenter effects in extrasensory perception. *Journal of the American Society for Psychical Research*, **69**, 135-149.

238-243 KANTHAMANI, H., & RAO, H. H. (1975). Response tendencies and stimulus structure. *Journal of Parapsychology*, **39**, 97-105.

244 LEVIN, J. A. (1975). A series of psi experiments with gerbils. *Journal of Parapsychology*, **39**, 363-365.

245-248 NEVILLE, R. C. (1975). Some aspects of precognition testing. *Research in Parapsychology*, 29-31.

249-251 DAVIS, J. W., & HAIGHT, J. (1976). Psi experiments with rats. *Journal of Parapsychology*, **40**, 54-55.

252-256 JACOBS, J., & BREEDERVELD, H. (1976). Possible influences of birth order on ESP ability. *Res Letter*, No. 7, 10-20.

257-263 DRUCKER, S. A., DREWES, A. A., & RUBIN, L. (1977). ESP in relation to cognitive development and IQ in young children. *Journal of the American Society for Psychical Research*, **71**, 289-298.

264-268 HARALDSSON, E. (1977). ESP and the defense mechanism test (DMT): A further validation. *European Journal of Parapsychology*, **2**, 104-114.

269-272 SARGENT, C. L. (1977). An experiment involving a novel precognition task. *Journal of Parapsychology*, **41**, 275-293.

273-276 BIERMAN, D. J. (1978). Testing the "advanced wave" hypothesis: An attempted replication. *European Journal of Parapsychology*, **2**, 206-212.

277-278 O'BRIEN, J. T. (1978). An examination of the checker effect. *Research in Parapsychology,* 153-155.

279 BRAUD, W. (1979). Project chicken little: A precognition experiment involving the SKYLAB space station. *European Journal of Parapsychology*, **3**, 149-165.

280-281 CLEMENS, D. B., & PHILLIPS, D. T. (1979). Further studies of precognition in mice. *Research in Parapsychology*, 156.

282 HARALDSSON, E., & JOHNSON, M. (1979). ESP and the defense mechanism test (DMT) Icelandic study No. III. A case of the experimenter effect? *European Journal of Parapsychology*, **3**, 11-20.

283-285 HARALDSSON, E. (1980). Scoring in a precognition test as a function of the frequency of reading on psychical phenomena and belief in ESP. *Res Letter*, No.10, 1-8.

286 SARGENT, C., & HARLEY, T. A. (1981). Three studies using a psi-predictive trait variable questionnaire. *Journal of Parapsychology*, **45**, 199-214.

287-290 THALBOURNE, M., BELOFF, J., & DELANOY, D. (1981). A test for the "extraverted sheep versus introverted goats" hypothesis. *Research in Parapsychology*, 155-156.

291-292 WINKELMAN, M. (1981). The effect of formal education on extrasensory abilities: The Ozolco study. *Journal of Parapsychology*, **45**, 321-336.

293 NASH, C. B. (1982). ESP of present and future targets. *Journal of the Society for Psychical Research*, **51**, 374-377.

294 SCHWARTZ, S. A., & DEMATTEI, R. J. (1982).. The Mobius psi-Q test: Preliminary findings. *Research in Parapsychology*, 103-105.

295-296 CRANDALL, J. E., & HITE, D. D. (1983).. Psi-missing and displacement: Evidence for improperly focused psi? *Journal of the American Society for Psychical Research*, **77**, 209-228.

297 TEDDER, W. (1983). Computer-based long distance ESP: An exploratory examination (RB/PS). *Research in Parapsychology*, 100-101.

298 JOHNSON, M., & HARALDSSON, E. (1984). The defense mechanism test as a predictor of ESP scores. Icelandic studies IV and V. *Journal of Parapsychology*, **48**, 185-200.

299 HARALDSSON, E., & JOHNSON, M. (1985). The defense mechanism test (DMT) as a predictor of ESP performance: Icelandic studies VI and VII. *Research in Parapsychology*, 43-44.

300-303 HESELTINE, G. L. (1985). PK success during structured and nonstructured RNG operation. *Journal of Parapsychology*, **49**, 155-163.

304-308 VASSY, Z. (1986). Experimental study of complexity dependence in precognition. *Journal of Parapsychology*, **50**, 235-270.

309 HONORTON, C. (1987). Precognition and real-time ESP performance in a computer task with an exceptional subject. *Journal of Parapsychology*, **51**, 291-320.

January 31, 1989

Ms. Jean V. Smith
Contracting Officer
Department of the Army
U.S. Army Medical Research And Development Command
Ft. Detrick, Frederick, Maryland 21701-5014

Dear Ms. Smith:

Enclosed please find one copy of the Project 1291 Final Report, Objective I, Task 1, "Meta-Analysis Of Forced-Choice Precognition Experiments." You now have one copy of each of the following nine FY 1988 report deliverables:

Project 1291, Contract Number DAMD 17-83-C-3106 FY 1988 Report Deliverables

Objective A, Task 3, Final Technical Report, "Enhanced Human Performance Investigation"

Annual Administrative Report, "Enhanced Human Performance Investigation"

Objective B, Task 1, Final Report, "Mass Screening For Psychoenergetic Talent Using A Remote Viewing Task"

Objective D, Task 1, Final Report, "Neurophysiological Correlates To Remote Viewing"

Objective E, Tasks 1 and 2, Final Report, "Feedback And Target Dependencies In Remote Viewing Experiments"

Objective E, Task 3, Final Report, "The Effects Of Hypnosis On Remote Viewing Quality"

Objective E, Task 4, Final Report, "Forced-choice Remote Viewing"

Objective F, Task 1, Final Report, "Applications Of Fuzzy Sets To Remote Viewing Analysis"

Objective I, Task 1, Final Report, "Meta-Analysis Of Forced-Choice Precognition Experiments"

Sincerely,

Catherine A. Flowers
Projects Coordinator

Enclosure

cc: Dr. Murray J. Baron
Mr. James O. Dolen
Dr. Edwin C. May
[redacted] SG1J

SRI International
333 Ravenswood Ave. • Menlo Park, CA 94025 • (415) 326-6200 • TWX: 910-373-2046 • Telex: 334486 • Facsimile: (415) 326-5512

Approved For Release 2000/08/08 : CIA-RDP96-00789R002200410001-2