

http://www.psiexplorer.com/blackmore_critique.htm

A Critical Examination of the Blackmore Psi Experiments

RICK E. BERGER
Science Unlimited Research Foundation
San Antonio, Texas

The Journal of the American Society for Psychical research
Vol 83, April 1989, 123-144

ABSTRACT: A critical examination of Susan Blackmore's psi experiment database was undertaken to assess the claims of consistent "no ESP" across these studies. Many inconsistencies in the experimental reports were found, and their serious consequences are discussed. Discrepancies were found between the unpublished experimental reports and their published counterparts. "Flaws" were invoked to dismiss significant results while other flaws were ignored when studies produced nonsignificant results. Experiments that were admittedly flawed in the unpublished reports were mixed with supposedly unflawed studies and published without segregation, creating the impression of methodological soundness. Two instances in which study chronology was reordered were found. Overall, it is concluded that Blackmore's claims that her database shows no evidence of psi are unfounded, because the vast majority of her studies were carelessly designed, executed, and reported, and, in Blackmore's own assessment, individually flawed. As such, no conclusions should be drawn from this database.

INTRODUCTION.

In early 1987, I was asked to review Susan Blackmore's (1986) autobiography, *The Adventures of a Parapsychologist* (Berger, 1988), in which she repeatedly claims that there was no sign of ESP in all of her experimental work. Questioning these claims, I amassed all of her publications that tested a psi hypothesis and, from these publications, produced a draft manuscript of meta-analyses of the Blackmore ESP experiments that suggested there might indeed have been psi effects in the database. Shortly after writing the draft, I procured a copy of Blackmore's unpublished doctoral dissertation—the original source material for the subsequent publications. Comparison of the dissertation and the later published reports revealed that my analyses, based only on her published reports, were inaccurate, as the published reports often did not veridically reflect the original data. My review of this work suggests that (a) working only from the published reports would inaccurately represent the original findings, and (b) reconciling the discrepancies of the later published papers with the unpublished dissertation and formally assessing the flaws in such studies must precede any formal meta-analysis of the Blackmore ESP experiments.

In a number of publications, Blackmore (cf. 1980a, 1980c, 1980d, 1981a, 1981b, 1983a, 1984, 1985a, 1985b, 1986, 1987, 1988) claims to have become increasingly skeptical about the existence of psi phenomena after "ten years of negative research in parapsychology" (Blackmore, 1987). Having been steeped in occult literature and practice, she entered the field of parapsychology as a fervent believer in the possibility of psi phenomena (Blackmore, 1986). In her writings, which span nearly a decade, she presents herself as an open-minded scientist. However, following the failure of her "very first experiment," she recorded in her diary: "I concluded that parapsychology is all a lot of rubbish and I should do something else!" (Blackmore, 1986, p. 35). Having reached this

conclusion, she continued to perform psi experiments for the duration of her doctoral program and earned a Ph.D. in parapsychology (in January 1980).

Blackmore's recent descriptions (e.g., Blackmore, 1985a, 1986, 1987) of her earlier research convince the reader that these experiments were scrupulously conducted and reported. One parapsychologist reviewing her autobiography (Blackmore, 1986) concluded:

For three years she carried out technically correct experiments intended to investigate ESP in relation to memory, ESP in small children, and ESP in the Ganzfeld condition. Except perhaps once in a preliminary experiment, she never obtained statistically significant evidence for the occurrence of a psi effect. (McConnell, 1987, p. 1)

A reviewer skeptical of ESP reached this conclusion:

With growing methodological rigour, she consistently fails to find any evidence to support her belief in the paranormal. . . . She gradually came to understand that ESP, telekinesis, Tarot card readings and the whole shabby collection of spurious contacts with a deeper reality that make up parapsychology are born of a failure to grapple with the cruel demands of decent scientific method. (Blinkhorn, 1987, p. 670)

In her autobiography, Blackmore recounts the comfort offered by her husband when she lamented her failure to obtain psi in experiments in which other researchers had succeeded: “Maybe they’re wrong and you are right. Maybe they haven’t done their experiments as carefully as you have” (Blackmore, 1986, p. 55).

Blackmore's statements concerning the lack of evidence for psi phenomena in general (cf. Blackmore, 1985a, 1987), the claims that her own research was consistently devoid of evidence for psi, and my review of her autobiography (Berger, 1988) prompted my examination of the database upon which her conclusions were drawn.¹ Specifically, the questions addressed were: Is her database sound? And, do the results support her claim of “no apparent psi effects” as she insists?

THE DATABASE BROADLY VIEWED.

In partial fulfillment of requirements for her doctoral dissertation, Blackmore reported 29 experiments conducted between October 18, 1976 and December 1978 (Blackmore, 1980c, pp. 135—136), of which 21 were eventually published as separate experiments in five peer-refereed parapsychology journal papers (see [Table 1](#)).² The experiments reported in the dissertation are “the results of all experiments carried out since October 1976” (Blackmore, 1980c, p. 133). This included many preliminary experiments and some very small studies which it may be thought do not warrant inclusion. The reason is to avoid any possibility of biased or selective reporting of results which could lead to a distortion of the overall picture. The only exceptions to this rule are some experiments which were carried out purely for the students' interest and from which systematic data were not recorded. (p. 133)

Blackmore described the care, preparation, and data analysis involved in the dissertation experiments:

My first experiments were far from perfect, but at least I did them. Sunday I spent frantically preparing my experiments to carry out on Monday. On those long Sunday evenings my friend Kim and I drew and redrew target pictures, sealed numbered lists in envelopes, tossed dice, stuck pins in random-number tables, and typed out questionnaires and answer sheets—all to be ready for Monday. . . . Yet, somehow I managed to analyze the results of the experiments as I did them, often staying up into the early hours of the morning with heaps of answer sheets and my trusty calculator, because I didn't want to waste the opportunity to test so many subjects. So I kept it up—one experiment a week for twenty weeks. And, as it turned out, for several years. (Blackmore, 1986, p. 32)

Following the dissertation research, Blackmore's publications focused on research on out-of-body experiences (OBEs), personality factors and belief in psi, and criticism of parapsychological

research. She has explicitly claimed that OBEs are not “paranormal” (Blackmore, 1986).³ Her dismissal of OBEs as subjective (nonpsi) experiences (as they well may be)⁴ earned her attention and praise from the skeptical community. A noted skeptic regarded Blackmore’s (1982) book on OBEs as

an excellent work, a book that earns my great respect for Dr. Blackmore’s abilities as a critical investigator in the best scientific tradition. . . . If there were more psychical researchers with the talents of Dr. Blackmore, the gulf that isolates psychical research from mainstream science could rapidly be bridged. (Alcock, 1983, p. 77)

Though she received no coauthorship for her work, Blackmore acted as the remote experimenter in a psi experiment for the “Bristol Series,” using Dick Bierman’s computer psi-testing software and her own baby as a subject in an attempted replication (Bierman, 1985b). The results were statistically significant and suggested a possible psi effect by her child. Troscianko and Blackmore (1985) later argued that the results may have been due to an artifact. Bierman (1985a) argued that the supposed artifact could not have accounted for the significant outcome in the original experiment.

Following the publication of the dissertation experiments, only one experiment (testing a psi hypothesis in which Blackmore’s name appears as an author) can be found in a refereed journal (Blackmore & Troscianko, 1985). This paper appeared in the *British Journal of Psychology* and was based, in part, on an experiment reported at the 1982 convention of the Parapsychological Association (Troscianko & Blackmore, 1983).

“Ten Years” of Negative Research.

The primary implication of Blackmore’s recent skeptical publications is that her “ten years of negative research” (see Blackmore, 1987) is a sound basis upon which she may conclude and promote the notion that parapsychology should be redefined as “a new psychical research—one without psi” (Blackmore, 1988, p. 58). Yet she says:

Impartiality forced me to admit that there is evidence for psi. It cannot all be successfully debunked, and there will always be more “successes” coming along. But I could not be impartial. The positive findings were other people’s and the negative ones were my own. So what could I do? (Blackmore, 1985a, p. 438)

She maintains that although she acknowledges the apparent replicability of research within laboratories elsewhere (cf. Blackmore, 1985b, p. 189; Blackmore, 1986, p. 97; Blackmore, 1987, p. 250), her personal experience has compelled her to disbelief.

A comparison of the experimental chronology from her dissertation (Blackmore, 1980c, pp. 135—136) and details from her autobiography (Blackmore, 1986) indicates that the bulk of her experimental psi research efforts (her dissertation experiments) occurred during a 2-year period (October 1976—December 1978), and that well before the end of this period she was a complete skeptic regarding psi phenomena (cf. Blackmore, 1987, p. 249).⁵

Conversion to Skepticism.

Blackmore (1987) helps pinpoint the time frame of her conversion from believer to skeptic. Though she was quick to pronounce parapsychology as “rubbish” following her first “failure” to confirm her (arbitrary) psi hypothesis (Blackmore, 1986, p. 35), her total conversion to skepticism apparently came after her series of three Tarot experiments (reported in Blackmore, 1983a). The first Tarot experiment produced significant results. Blackmore states that after the last Tarot experiment (completed in November 1978), she “chose this point to say, ‘I think that, however many more experiments I do on psi, I am probably not going to find it’ (Blackmore, 1987, p. 249).

In describing the “cognitive dissonance”⁶ she has experienced as a result of her failure to find evidence of psi, she has stated:

I found myself simply not believing in psi anymore. I really had become a disbeliever. Like one of those doors with a heavy spring that keeps it closed, my mind seemed to have changed from closed belief to closed disbelief. (p. 249)

Apparently, by the time she had received her degree (in 1980), she was a confirmed skeptic regarding psi.⁷ Blackmore has stated: “If the experimenter’s beliefs or expectations play a role [in the experiments], then the later experiments never stood a chance” (Blackmore, 1983b, p. 17). These later experiments included her oft-mentioned, but unpublished, Ganzfeld study.

OVERVIEW OF THE DATABASE.

The “Notes on Experimental Section” in Blackmore’s dissertation reveals that of the reported experiments, 12 were “carried out without optimum methods and for exploratory purposes” (Blackmore, 1980c, p. 133). From the remaining studies, five reports in refereed publications encompassing 21 experiments emerged from the dissertation research in parapsychological journals (Blackmore, 1980a, 1980d, 1981a, 1981b, 1983a). Most of the dissertation experiments were group experiments conducted in single classroom sessions using her students as subjects.

A review of these publications revealed a number of discrepancies between the original studies (as reported in the dissertation) and the later, published versions. Some of the discrepancies are outlined in the brief review of the publications below.⁸

Correlations Between ESP and Memory (Blackmore, 1980a, “Correlations”)

Six experiments were reported, two of which were labeled as “preliminary” and “without optimum methods” in the dissertation (Blackmore, 1980c). The ordering of experiments within this publication presents a false chronology of the sequence of these studies (detailed below in “Reordering of Published Experiments”).

ESP in Young Children (Blackmore, 1980d, “Children”)

Two experiments were conducted with small children as subjects. These experiments “were not replications of Spinelli’s work but drew heavily on his findings, using similar tasks and children of the age he had found best” (Blackmore, 1985a, p. 428). Whereas Spinelli had tested 1,000 subjects to achieve his reported results (Spinelli, 1977), Blackmore used 19 and 48 children in her two studies. Neither of Blackmore’s studies showed an overall psi effect.

The Effect of Variations in Target Material on ESP and Memory (Blackmore, 1981a, “Target”)

Four of the six experiments were labeled “preliminary” and “without optimum methods” in the dissertation. This paper also presents a false chronology of the sequence of studies (detailed below in “Reordering of Published Experiments”).

In the unpublished dissertation, Experiment 1 (in which Blackmore served as the sole subject) reports that “there were too few trials to conclude that there is no effect” and that “the results of this exploratory study are included only for the sake of completeness” (Blackmore, 1980c, p. 171, dissertation). In the published version (Blackmore, 1981a, “Target”), no such disclaimer is noted.

In the Results section of the publication it is noted that “there were more hits when visualising pictures” rather than words, “but not significantly so ($t = 2.74$, $df = 4$, $p = 0.52$)” (Blackmore, 1981a, p. 11, “Target”). The probability value should read $p = .052$; this value was reported

incorrectly only in the published version. It is also so close to significance as to perhaps deserve some comment.

In Experiment 4 of this series, the number of subjects is incorrectly reported in the article as 23 (the correct number is 28; see Blackmore, 1980c, p. 177, dissertation). Two “faults in the design” of Experiment 5 were reported but were, according to Blackmore, “unlikely to be responsible for the uniformly chance results obtained here” (Blackmore, 1981a, p. 19, “Target”). The last experiment reported in this series, labeled “Main Experiment” in the Research Letter but not so distinguished in the dissertation, used words as ESP targets. Some words were “common,” some “uncommon,” and some “naughty” (such as “sperm,” “penis,” and “screw”).

Errors and Confusions in ESP (Blackmore, 1981b, “Errors”)

Four experiments were presented, of which two were deemed “preliminary” and “without optimum methods” in the dissertation.

In the introduction to this four-experiment report, Blackmore stated:

Three pilot studies were carried out. They were performed quickly using large numbers of subjects with only one target order. They were therefore subject to a stacking effect. . . . Because these studies suffered from various flaws they are only described in outline here. (Blackmore, 1981b, pp. 54—55, “Errors”)

An examination of Pilot Study 1 (in Blackmore, 1981b, “Errors”) reveals that in this first experiment of the dissertation series significant results were found, albeit not in the condition that Blackmore’s theory favored. The fourth experiment, called “Main Study,” also produced a significant outcome. Both will be discussed further under “Invoking Study Quality When Outcome is Significant.”

Though Blackmore cautions that three of the four studies are flawed (pp. 54—55), she nevertheless later aggregated the studies in order to draw conclusions. She first states that “Experiment 1 may be excluded because of the faulty method used. This leaves two adequate experiments providing very different results” (Blackmore, 1981b, p. 65, “Errors”).⁹ In her Table 6, she compares the outcome of the two “adequate” experiments¹⁰ and concludes that “in neither was there any evidence of ESP occurring” (p. 65, italics added). This conclusion is drawn despite the fact that in the Main Study “a one-way ANOVA shows a significant effect of word type on the number of hits” (p. 63). Direct hits showed significant missing ($t = -3.14$, 58 df, $p = .003$).

Divination With Tarot Cards: An Empirical Study (Blackmore, 1983a, “Tarot”)

Three experiments were conducted using Tarot cards. Blackmore, having used Tarot cards for divination for many years, believed that they “worked” and that she “might find that psi manifested itself in the cards while being shy of laboratory experiments” (Blackmore, 1986, p. 61). Though I would argue that such a study fails to meet the minimum standards for a proper psi experiment (see, e.g., Chapter 4 by Morris in Edge Morris, Palmer, & Rush, 1986), and nonpsychic Blackmore herself served as the “psychic reader” for this experiment, it is considered by Blackmore to be among her “psi” experiments.¹¹

In the first study, Blackmore acted as the card reader for students whom she knew well. The study produced a significant positive outcome—also the largest absolute effect size in all of her dissertation experiments (see T1 in [Table 2](#)).

While Blackmore was reporting the significant results of this study to a meeting of the Cambridge University Society for Psychical Research, Carl Sargent pointed out to her that her statistical measure assumed independence of ratings and, as the subjects knew one another, ratings were not independent (Blackmore, 1986, pp. 66—67). Two “replications” of this study were completed and

reported: The second experiment changed the conditions of the first (strangers were used as opposed to friends), and the third study attempted Tarot readings by mail for strangers on a different continent. Though the changes in experimental conditions were considerable, when discussing this series as evidence against psi, her claims sound as though she performed literal replications: “So I repeated the experiment twice more with subjects who did not know one another. I expect you can predict the results I obtained—entirely nonsignificant” (Blackmore, 1987, p. 249).

Betty Markwick, a statistician who is highly regarded by the skeptical community for her exposure of the manipulation of the Soal-Goldney data, recently reanalyzed Blackmore’s Tarot experimental data (Markwick, 1988). Using statistical methods that are valid considering the design of the experiment, Markwick found that the data for the first experiment remain significant.

Ganzfeld Experiment (Unpublished)¹²

The experimental report of Blackmore’s Ganzfeld study (dated December 10—16, 1978 [Blackmore, 1980c, pp. 135—136, dissertation]) is found only in her unpublished dissertation. In a separate paper, Blackmore (1980b) published an evaluation of the “filedrawer” of unpublished Ganzfeld experiments. Studies were considered either “adequate” or “inadequate” “before accepting them as valid” (p. 213). Factors associated with inadequacy included the “use of picture targets without a duplicate set for judging” (p. 214).

The Blackmore Ganzfeld study was the last of the dissertation experiments. It contains numerous admitted flaws, including those described in the randomization procedure:

The agent [usually a friend or relation to the receiver] . . . shuffled the four envelopes to choose the target. This method, which obviously allows for both cheating and accidental non-randomness, was used for the first 20 sessions. Thereafter the target was chosen from random number tables prior to the experiment. (Blackmore, 1980c, p. 282)

A total of 36 sessions was conducted, but the better randomized sessions were “unplanned” and only conducted because “so many Ss were keen to have another session” (Blackmore, 1980c, p. 284). Further, only one target set was used (another fatal flaw according to Blackmore), and no independent judging was done. This inadequate Ganzfeld study is frequently cited among her failed psi experiments with the implication that it was methodologically sound (cf. Blackmore, 1986, pp. 99—107; 1987, pp. 247—248), despite the fact that Blackmore is aware that this experiment would not be accepted for publication in any peer-refereed parapsychological journal.

Blackmore and Troscianko (1985)

Three experiments are reported, but only Experiment 3 involved a psi task. In each experiment, subjects were classified as “sheep” or “goats.” The dichotomy of subjects into sheep and goats was performed by mean splits, with the actual mean only reported for Experiment 3. Hence, we have no frame of reference for judging whether ‘sheep’ in one experiment may have been classified as “goats” in another (or vice versa). Experiment 3 was a test for psychokinesis (PK), using a hardware random number generator (RNG) and computer psi task. The paper reports that there was no evidence of a deviation from chance scoring between “sheep” and “goats.”

SPECIFIC CRITICISMS

Reordering of Published Experiments

Only by examining the “Schedule of Experiments” in the unpublished dissertation (Blackmore, 1980c, pp. 135—136) and comparing this to the published versions can one reconstruct the actual sequence of experiments. [Table 2](#) shows the chronological order of the published dissertation psi experiments. The column “t (ESP main measure)” presents, where available, the reported t test for an ESP main effect. (Some studies focused on correlations of ESP scoring with a second measure and did not report ESP main effects.)

The first instance of reordering was found in the six experiments reported in Blackmore, 1980a (“Correlations”), that were conducted over a 2-year period (see Tables 1 and 2). Blackmore states in the introduction to her paper:

The results of 6 experiments are reported here. The preliminary experiments were carried out with the intention of finding the best methods to use in later experiments. Since they suffer from methodological weaknesses they are only reported in outline here (for further details see Blackmore, 1980~cI). (Blackmore, 1980a, p. 133, “Correlations”)

The experiments are reported as “Experiments 1—5” and “Main Study Experiment 6.” The actual chronological order (i.e., the order of study completion as reported in the dissertation) was 3, 1, 4, 2, 6, 5.

In the conclusion of “Main Study Experiment 6” in the journal publication, it is stated: “On the basis of the preliminary experiments several hypotheses were made and tested in a final experiment but were not confirmed” (Blackmore, 1980a, p. 143). The “final” experiment (completed, according to the dissertation chronology, on December 4, 1978) preceded the fifth experiment (completed December 11, 1978) by one week.

A second instance of reordering was found in the six experiments reported in Blackmore, 1981 a (“Target”), which were also conducted over a 2-year period (see Tables 1 and 2).¹³ The ordering of these experiments in the published version of this experimental series suggests that Experiments 1 and 2 logically and temporally preceded Experiments 3 through 6. Experiment 6 has been labeled “Main Experiment,” and the other five are labeled “Preliminary Experiments 1—5” in the refereed publication, though the dissertation chronology reveals that Experiments 3—6 actually predate Experiments 1 and 2 (e.g., Experiment 2 actually was carried out 2 years after Experiment 3 [Blackmore, 1980c, pp. 135—136, dissertation]).

In the introduction to “Main Experiment” (Experiment 6), Blackmore states: “Problems found in the previous experiments were eliminated and all the subjects had individual target orders” (Blackmore, 1981a, p. 19, “Target”) for this experiment. I believe that there is no other way to interpret this remark, except to believe that the “Main Experiment” followed the completion of the previous five experiments (incorporating knowledge gained from them) when in fact it had not. At most, the “Main Experiment” followed the completion of 3 of the reported experiments (3, 4, and 5). In Study 1, Blackmore served as the sole subject. As she had never claimed either spontaneous or laboratory evidence for psi ability, it is not surprising that this study showed a chance outcome. She then replicated the procedure in Study 2 using her students as subjects and found overall significant psi missing.

The rearrangement of study order obfuscates a substantial decline over the 2-year period of the main measures of ESP scoring (from above to below chance) that is apparent when the data are properly ordered ($r [4] = -.80$, $p = .056$).

Methodological Flaw Throughout Database

Most critics would consider as fatally “flawed” any psi study in which the data were scored by the subjects themselves. Three of the five publications that emerged from the dissertation research (1980a, “Correlations”; 1981a, “Target”; 1981b, “Errors”) were composed of experiments conducted during classroom sessions with students in her parapsychology courses. Of the 16 published experiments in these three publications, in most of them the procedure clearly states (in the dissertation) that the subjects scored all or part of the experimental data, usually by scoring the data of a neighboring student (see Table 1).

Though the description by Blackmore (1981b, “Errors”) is virtually verbatim from the dissertation version, the last paragraph of the procedure (for Pilot Study 1) has been omitted from the published version. The omitted paragraph includes: “When all Ss had completed the task they were asked to

give their answer sheets to a neighbour for checking” (Blackmore, 1980c, p. 140, dissertation). She further states in her dissertation: “In this experiment the Ss marked each others’ answer sheets. Obviously this introduces the possibility of cheating. . . . [T]his procedure was used in all experiments in the year 1976—7 (1—9 in schedule of experiments)” (p. 144).

Invoking Study Quality When Outcome is Significant

The invocation of flaws throughout Blackmore’s publications appears to be systematically related to study outcome. In instances where results were significant, and possibly indicative of psi, Blackmore dismisses the results as uninterpretable due to flaws or faults in experimental design. This can be seen in Blackmore, 1980a (“Correlations,” Experiment 2), 1981a (“Targets,” Main Experiment), 1981b (“Errors,” Pilot Study 1 and Main Study), and 1983a (“Tarot,” Experiment 1).

Significant effects that apparently supported her memory theory of psi (significantly more associative hits than expected, as well as significantly more associative than perceptual errors) were published as “Pilot Study 1” in Blackmore (1981b, “Errors”). In the discussion section of the dissertation, she states: “This may appear to support the hypothesis that errors made in ESP more closely resemble those made in memory than in perception” and that the results “appear to support the hypothesis that associative errors occur more frequently” (Blackmore, 1980c, p. 142). She cites numerous flaws in the study as reasons to dismiss the outcome. These include a stacking problem, target problems, and subjects scoring their own data (which Blackmore suggests may introduce the possibility of cheating).

When significant results were obtained in the “Main Study” of Blackmore, 1981b (“Errors”), she suggests several interpretations of the data and then claims:

However, none is universally accepted and I had not decided, prior to the experiment which model I intended to use. It therefore seems that no definite conclusions can be drawn from the results obtained. The results highlight the fact that possibly untenable assumptions were made in designing the experiment. (Blackmore, 1981b, p. 64, “Errors”)

In Blackmore (1981a, “Target”), study quality was not invoked to dismiss a significant result—instead the result was simply not reported. Here, the description of “Main Experiment” virtually reproduces the original dissertation report except for the following omission: “Or for ESP score $r = 0.286$ ($z = 2.0$ $p = 0.045^*$). This correlation is significant but is in the direction opposite to that predicted by the negative response bias hypothesis” (Blackmore, 1980c, p. 185, dissertation).

In Blackmore’s dissertation, the discussion states that the experiment (later published as Blackmore, 1981a, “Target”) “was poorly designed” (1980c, p. 185). Both the significant result and reference to the study’s poor design have been omitted in the published version.

Blackmore’s first Tarot experiment’s significant outcome (Blackmore, 1983a, “Tarot”) was dismissed on two grounds: First, the significance “depends on the use of 1-tailed tests” (p. 98). Despite the fact that the tests were planned to be one-tailed and “that differences in the opposite direction would be meaningless” (p. 99), Blackmore then says that “it could be argued that 2-tailed tests should always be used in parapsychological experiments because of the difficulty of predicting scoring directions” (p. 99). Blackmore’s second flaw in this study was the statistical problem mentioned earlier, though Markwick’s recent (1988) reanalysis suggests that the results remain significant with proper statistical evaluation.

Ignoring Study Quality When Outcome is Nonsignificant

Throughout the dissertation, Blackmore acknowledges that individual studies are flawed in many ways. In a majority of the published experiments (see [Table 1](#)), Blackmore acknowledges certain experimental flaws, yet when conclusions based on the experiments are made, experiments with “flawed” designs are weighted the same as experiments that had “proper” designs.

In the original description of the experiment later reported as Experiment 5 in Blackmore (1981a,

“Target”), she comments:

There were two major faults in the design of this experiment. Firstly the same target set was used for all Ss. . . . Secondly word length was confounded with target type. . . . A final fault was that the design of the experiment made checking extremely difficult and laborious, so increasing the possibility of errors. . . . These faults, however, might be expected to produce spurious differences, but are unlikely to be responsible for the uniformly chance results obtained here. (Blackmore, 1980c, p. 181)

It appears that it is Blackmore’s argument that flaws can plausibly only lead to false positives (Type I errors). It is beyond the scope of this paper to elaborate, but there are a number of design flaws that can lead to false negatives (Type II errors). These include, but are not limited to, inadequate sample size (low statistical power), weak or inappropriate statistical tests, sampling from inappropriate populations, experimenter expectancy effects, demand characteristics, and the faulty operationalization of dependent measures.

Many skeptics, when appraising positive evidence for psi, consider flaws of any sort as evidence of a “dirty test tube” (e.g., Hyman, 1985a, pp. 41—42). The gist of the dirty test tube argument is that such flaws can be regarded as “symptoms” and that this “suggests a casualness that is inappropriate for an investigation that is being asked to carry part of the burden for asserting the existence of phenomena that many scientists find difficult to believe” (Hyman, 1985a, p. 84). One must hold to the same standards of experimental design in any parapsychological study, regardless of its outcome. Some skeptics, including Blackmore, argue that differing standards of experimental design can be held depending on study outcome: Significant positive outcomes must have tighter designs than the same study with a negative outcome. This post hoc determination of experimental criticism leads to the paradox exemplified by the Blackmore work: Had such work produced consistently positive outcomes, the results could all be dismissed as having arisen from design flaws and the “dirty test tube.” Because the studies did not yield consistently positive results, the flaws can be overlooked and the database viewed as a coherent body of evidence that converges on the conclusion that psi does not exist. Negative conclusions based on flawed experiments must not be given more weight than positive conclusions based on the same flawed experiments. The meaningfulness of a scientific study is determined by how well the dependent measure was operationalized, not by whether the experimental result fits one’s preconceptions of what the outcome “should have been.”

Misreporting the Original Data

Blackmore (1986), arguing that she couldn’t study the psi process because she never found any ESP, stated: “At one point I calculated that I had performed thirty-four independent [*italics added*] significance tests and just two were significant—remarkably close to chance expectation” (p. 53). This claim was repeated elsewhere, almost verbatim, at a 1983 conference sponsored by the Parapsychology Foundation (Blackmore, 1985b, p. 188). It is found, in a modified form, in a chapter in *A Skeptic’s Handbook of Parapsychology*: “At one point I calculated that I had performed 34 independent [*italics added*] significance tests in almost as many experiments and obtained two values significant at the 0.05 level” (Blackmore, 1985a, p. 427).

The original published data supporting this claim can be found in Blackmore (1980a, “Correlations”), which states: “If all analyses are considered (though not all are independent) [*italics added*] in a total of 34 significance tests 2 were significant at $<.05$ ” (p. 145). If one traces this published quote back to its original data as reported in the unpublished doctoral dissertation, one finds that the 6 studies reported in Blackmore (1980a, “Correlations”) were originally reported in dissertation Chapter 8 (“Correlations Between ESP and Memory Ability”) in which 8 experiments are reported.¹⁴ Reviewing the 8 experiments in Chapter 8 of the dissertation, Blackmore concludes:

Of 12 correlation coefficients reported only one is significant. . . . In fact many other correlation coefficients were reported and significance tests carried out. If all are included (though all are not independent) [*italics added*] in a total of 34 significance tests 2 were significant (p. 216).¹⁵

Thus, in the retelling, significance tests that were originally nonindependent and obtained from a series of 8 experiments were later reported to a skeptical audience as being independent and derived from almost 34 experiments.

In her autobiography, Blackmore describes the following experiment, and then calls it her “very first experiment” which “launched [her] into the beginnings of a quandery [sic] which took [her] more than ten years to resolve” (Blackmore, 1986, P. 34). She describes this experiment as follows:

In each [target] set the key picture (a caterpillar, for example) resembled two other pictures; one (a butterfly) was closely associated with it, while the other (a train) looked very similar but was not associated with it. In the actual experiment the target picture (the caterpillar) was sealed in an envelope and hidden from the subjects. All the subjects, my hundred [italics added] students, sat in their class with a sheet containing many pictures. They had to choose which one they thought was the target. . . . If the students had picked the caterpillar most often, with train in second place, I would have had to conclude that perhaps the perceptual theory was better than my memory theory. . . . They might have picked the caterpillar most frequently but not picked the train or the butterfly more often than all the other unrelated pictures. . . . But what happened was none of these. There were simply no meaningful results at all. Neither train nor butterfly [italics added] (and their equivalents in the other set of pictures), nor even the caterpillar, the actual target, was systematically picked more often than one would expect by chance [italics added]. In other words, there was no sign of any ESP. (Blackmore, 1986, pp. 34—35)

The published account of her ‘very first experiment’ (which is, according to her dissertation chronology, reported as “Pilot Study 1” in Blackmore, 1981b) states:

There are significantly more type 2 (associative) errors than expected ($t = 3.48$; $df = 5$; $p = <.04$).¹⁶ In addition for the key pictures only, a direct comparison can be made and this shows that there were significantly more type 2 (associative) than type 3 (perceptual) errors. This may appear to support the hypothesis that errors made in ESP more closely resemble those made in memory than in perception. (Blackmore, 1981b, p. 56. “Errors”)

The results of this experiment are then dismissed entirely because “inadequacies in the experimental design make such a conclusion unwarranted” (Blackmore, 1981a, P. 56, “Errors”).¹⁷ (One week later, a second experiment failed to replicate the results of the first experiment.) Following one or the other of these experiments, Blackmore recorded in her diary that “parapsychology is all a lot of rubbish” (Blackmore, 1986, p. 35).

Blackmore seems to be arguing that a flawed study with a significant outcome is equal to a negative outcome. To claim that “neither train nor butterfly was systematically picked more often than one would expect by chance” and that “there was no sign of any ESP” contradicts the results from the first experiment. These results have apparently been dismissed due to the failure to achieve perfect replication in the second attempt.

Possible Psi Effects in Database

During my aborted meta-analyses of Blackmore’s published work, I was struck by patterns in the data suggestive of the operation of psi.¹⁸ Much of the veracity of the published work is now in question, when compared with its original unpublished source. Without a serious meta-analysis on the original unpublished source material, complete with weighting for flaws (which can plausibly be shown to relate to study outcome), the issue of whether the Blackmore experiments show evidence for psi cannot be resolved. As evidenced by the recent Hyman/Honorton exchanges regarding the meta-analyses of the Ganzfeld research (Honorton, 1985; Hyman, 1985b), such an approach cannot resolve the integrity of a database—it can only point out its weaknesses and make recommendations for future research. Combining the results across the Blackmore database of experiments would certainly yield heated disagreement if positive results emerged, though the negative conclusions drawn by Blackmore about each published experimental series and their combined results have remained, until now, unchallenged.

CONCLUSIONS

After some period of time spent in attempting to become “a famous parapsychologist” (Blackmore,

1986, p. 163) and believing that she had failed to do so, Blackmore's attitude toward the reality of psi moved from "closed belief to closed disbelief" (Blackmore, 1987, p. 249). Though this attitude change is suggested to have been abrupt, as in the previous quote, it actually appears to have been a very gradual process, exacerbated by a number of factors (Berger, 1988). Whether the dissertation experiments that were concomitant with her increasingly skeptical belief system were "fair tests" of psi cannot be determined. We can, however, assess the integrity of the database as reflected by the original unpublished dissertation, subsequent partial publications from it, and Blackmore's polemical works that refer to this database.

Much of Blackmore's work is considered flawed by her own self-assessment. Serious discrepancies were found between the unpublished dissertation experiments and subsequent published journal reports. The claim of "ten years of psi research" actually represents a series of hastily constructed, executed, and reported studies that were primarily conducted during a 2-year period. Prior to the end of this period, she had moved to "closed disbelief." Her other "research" consists primarily of informal hypothesis testing and cursory examination of areas that do not (or may not) directly assess the psi hypothesis at all (e.g., mystical experiences, ghosts, poltergeists, out-of-body experiences, near-death experiences, and apparitions). She has admitted that she "assumed that all these odd and inexplicable things . . . were related and that one explanation would do for all" (Blackmore, 1987, p. 245). Though she is loath to publicly state that psi phenomena do not exist, she has made a career of promoting the idea that parapsychology should be redefined to exclude the psi hypothesis (see, e.g., Blackmore, 1985a, 1985b, 1988).¹⁹

For any conclusions to be drawn regarding the presence or absence of psi effects in her database, a serious meta-analysis with weighting of each study for flaws would be necessary. That many of the studies in this database may have insufficient statistical power to detect small effects and were not designed with sufficient intention to optimize the detection of psi can only serve to bias any informal meta-analysis toward a nonsignificant outcome.

Research into "experimenter expectancy" effects and "demand characteristics" suggests that, from a social psychological perspective, she may have influenced her subjects to perform in a manner consistent with her "no psi" hypothesis. Even if such studies had yielded significance, it is clear that such outcomes by now would have been scrutinized and dismissed by skeptics and proponents alike because of their experimental flaws and the haphazard conceptualization and execution of these studies.

Meanwhile, Blackmore is extremely vocal in decrying psi research in her writings, on television and radio, and before the skeptical advocacy group CSICOP (the Committee for Scientific Investigation of Claims of the Paranormal), citing her own work as the basis for her strong convictions.²⁰ Her recent polemical works often seriously misrepresent her original work, with the distorted information being more consistent with her current skeptical world view. The present overview of her database suggests that drawing any conclusions, positive or negative, about the reality of psi that are based on the Blackmore psi experiments must be considered unwarranted.

REFERENCES

- ALCOCK, J. E. (1983). Psychology of the out-of-body experience. *Skeptical Inquirer*, 8, 74—77.
- BERGER, R. E. (1988). Review of *The Adventures of a Parapsychologist* by S. Blackmore. *Journal of the American Society for Psychical Research*, 82, 374—384.
- BERGER, R. E. (in preparation). *Experimental Flaws and the Skeptics' Double Standard*.
- BIERMAN, D. J. (1985a). An impossible artifact. *European Journal of Parapsychology*, 6, 99—103.

- BIERMAN, D. J. (1985b). A retro and direct PK test for babies with the manipulation of feedback: A first trial of independent replication using software exchange. *European Journal of Parapsychology*, 5, 373—390.
- BLACKMORE, S. J. (1980a). Correlations between ESP and memory. *European Journal of Parapsychology*, 3, 127—147.
- BLACKMORE, S. [J.] (1980b). The extent of selective reporting of ESP ganzfeld studies. *European Journal of Parapsychology*, 3, 213—219.
- BLACKMORE, S. J. (1980c). *Extrasensory Perception as a Cognitive Process*. Unpublished doctoral dissertation, University of Surrey, Guildford, England.
- BLACKMORE, S. [J.] (1980d). A study of memory and ESP in young children. *Journal of the Society for Psychical Research*, 50, 501—520.
- BLACKMORE, S. J. (1981a). The effect of variations in target material on ESP and memory. *Research Letter*, 11, 1—26.
- BLACKMORE, S. J. (1981b). Errors and confusions in ESP. *European Journal of Parapsychology*, 4, 49—70.
- BLACKMORE, S. J. (1982). *Beyond the Body: An Investigation of Out-of-the-Body Experiences*. London: Heinemann.
- BLACKMORE, S. J. (1983a). Divination with Tarot cards: An empirical study. *Journal of the Society for Psychical Research*, 52, 97—101.
- BLACKMORE, S. J. (1983b). Prospects for a psi-inhibitory experimenter [Summary]. In W. G. Roll, J. Beloff, & R. A. White (Eds.), *Research in Parapsychology 1982* (pp. 17—20). Metuchen, NJ: Scarecrow Press.
- BLACKMORE, S. [J.] (1984). ESP in young children: A critique of the Spinelli evidence. *Journal of the Society for Psychical Research*, 52, 311—315.
- BLACKMORE, S. [J.] (1985a) The adventures of a psi-inhibitory experimenter. In P. Kurtz (Ed.), *A Skeptic's Handbook of Parapsychology* (pp. 425—448). Buffalo, NY: Prometheus Books.
- BLACKMORE, S. J. (1985b). Unrepeatability: Parapsychology's only finding. In B. Shapin & L. Coly (Eds.), *The Repeatability Problem in Parapsychology* (pp. 183—206). New York: Parapsychology Foundation.
- BLACKMORE, S. [J.] (1986). *The Adventures of a Parapsychologist*. Buffalo, NY: Prometheus Books.
- BLACKMORE, S. J. (1987). The elusive open mind: Ten years of negative research in parapsychology. *Skeptical Inquirer*, II, 244—255.
- BLACKMORE, S. [J.] (1988). Do we need a new psychical research? *Journal of the Society for Psychical Research*, 55, 49—59.
- BLACKMORE, S. [J.], & TROSCIANKO, T. (1985). Belief in the paranormal: Probability judgments, illusory control, and the “chance baseline shift.” *British Journal of Psychology*, 76, 459—468.

BLINKHORN, S. (1987). One knock for “no.” *Nature*, 325, 670—671. EDGE, H. L., MORRIS, R. L., PALMER, J., & RUSH, J. H. (1986). *Foundations of Parapsychology: Exploring the Boundaries of Human Capability*. Boston, MA: Routledge & Kegan Paul.

HONORTON, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology*, 49, 51—91.

HYMAN, R. (1985a). A critical historical overview of parapsychology. In P. Kurtz (Ed.), *A Skeptic's Handbook of Parapsychology* (pp. 3—96). Buffalo, NY: Prometheus Books.

HYMAN, R. (1985b). The ganzfeld psi experiment: A critical reappraisal. *Journal of Parapsychology*, 49, 3—49.

MARKWICK, B. (1988). Re-analysis of some free-response data. *Journal of the Society for Psychical Research*, 55, 220—222.

MCCONNELL, R. A. (1987). Left brain skepticism: A review of Dr. Susan Blackmore's *Adventures of a Parapsychologist*. Unpublished manuscript.

SHAPIN, B., & COLY, E. (EDS.). (1985). *The Repeatability Problem in Parapsychology*. New York: Parapsychology Foundation.

SPINELLI, E. (1977). The effects of chronological age on GESP ability [Summary]. In J. D. Morris, W. G. Roll & R. L. Morris (Eds.), *Research in Parapsychology 1976* (pp. 122—124). Metuchen, NJ: Scarecrow Press.

TROSCIANKO, T., & BLACKMORE, S. J. (1983). Sheep-goat effect and the illusion of control [Summary]. In W. G. Roll, J. Beloff, & R. A. White (Eds.), *Research in Parapsychology 1982* (pp. 202—203). Metuchen, NJ: Scarecrow Press.

TROSCIANKO, T., & BLACKMORE, S. J. (1985). A possible artifact in a PK test for babies. *European Journal of Parapsychology*, 6, 95—97.

TRUZZI, M. (1987). Zetetic ruminations on skepticism and anomalies in science. *Zetetic Scholar*, No. 12—13, 7—20.

Table 1
ALL BLACKMORE PUBLISHED DISSERTATION PSI EXPERIMENTS

Date Published (Date: study number)	Date Performed*	Flaws Identified by Blackmore		Sig. Results Noted in Report?	Reviewer Comments (see Key)
		In Dissertation (see Key)	In Journal (see Key)		
"Correlations"					
1980a:1	02/28/77	1,2,3,4	2,3	no	
1980a:2	12/05/77	3,4	3	yes	
1980a:3	02/07/77	1,2,3,4	2,3	no	
1980a:4	11/21/77	4		yes	
1980a:5	12/11/78	4		no	
1980a:6	12/04/78	4	4	no	12,15
"Children"					
1980d:1	06/30/78	17	17	no	13
1980d:2	07/28/78			no	13
"Target"					
1981a:1	05/11/78	1,18,19	14,18	no	
1981a:2	11/29/78	4	4	yes	20
1981a:3	11/22/76	1,2,4	2,4	no	
1981a:4	02/14/77	1,2,4		no	21
1981a:5	11/29/77	1,2,5,6	2,5,6	no	
1981a:6	01/30/78	4,7	4	yes	12,15,22
"Errors"					
1981b:1	10/18/76	1,2,4,8,9	2,8,9,10	yes	
1981b:2	10/25/76	1,2,4,8	2	no	
1981b:3	11/14/77	2,4	2	no	
1981b:4	11/29/77	4	4	yes	12
"Tarot"					
1983a:1	03/??/78	11	11	yes	16,23
1983a:2	06/??/78			no	16
1983a:3	11/??/78			no	16

* All dates reflect the date of completion of study from "Schedule of Experiments" (Blackmore, 1980c, pp. 135–136).

[back](#)

Table 2
BLACKMORE PUBLISHED PSI EXPERIMENTS ARRANGED IN CHRONOLOGICAL ORDER (BASED
ON "SCHEDULE OF EXPERIMENTS" [BLACKMORE, 1980C, P. 135, DISSERTATION])

Exp. Number in 1980c	Experimental Order in Publication ^a					Date Mn, Dy, Yr	<i>t</i> (ESP main measure)	<i>df</i>	Effect Size (Cohen's <i>d</i>) ^b
	8 0 a	8 0 d	8 1 a	8 1 b	8 3 a				
1				1		10/18/76	-0.99	8	-0.330
2				2		10/25/76	+1.25	8	+0.417
5(a)			3			11/22/76	+0.15	45	+0.022
8,3,4	3					02/07/77	+1.09	22	+0.227
5(b)			4			02/14/77	-0.88	27	-0.166
3,4,7	1					02/28/77	+1.12	19	+0.250
D				3		11/14/77	n/a	n/a	n/a
E,C,F	4					11/21/77	+2.35	49	+0.332
6			5			11/29/77	+0.04	55	+0.005
G(a)				4		11/29/77	-3.14	58	-0.409
H,C,F	2					12/05/77	+1.98	35	+0.330
J			6			01/30/78	-0.76	44	-0.113
T1				1		03/??/78	+3.54	9	+1.119
K			1			05/11/78	-0.92	9	-0.291
T2				2		06/??/78	+0.82	18	+0.188
L		1				06/30/78	-0.15	39	-0.024
M		2				07/28/78	+0.21	52	+0.029
T3				3		11/??/78	n/a	n/a	n/a
16			2			11/29/78	-2.13	49	-0.301
17	6					12/04/78	-0.32	31	-0.057
18,14	5					12/11/78	-0.53	19	-0.119

Key: 1980a: "Correlations" 1980c: dissertation 1980d: "Children"

1981a: "Target" 1981b: "Errors" 1983a: "Tarot"

^a Numbers represent chronological order of experiments as presented in journal publication.

^b Calculated using the formula:

$$d = \frac{t}{\sqrt{n}}$$

[back](#)

Key to Table Comments

¹ Blackmore (1980c) cautioned that "the term 'preliminary' is used loosely to apply to those experiments which were carried out without optimum methods and for exploratory purposes. This refers particularly to experiments 1—9 carried out in 1976—7, and experiment G part 2 and K" (p. 133).

² Study used a single target order, allowing possible stacking effect (Blackmore, 1980c, p. 175).

³ ESP tests were conducted prior to the memory tests and Ss already knew their ESP scores when they took the memory test. Conceivably this could lead to a spurious correlation between the two" (Blackmore, 1980c, p. 193).

⁴ Subjects scored all or part of the experimental data. Blackmore states: "It was thought that the subjects should have feedback on their scores as soon as possible after the tests so as to maintain their interest. For this reason they were allowed to mark each others' answer sheets. This necessarily introduced the possibility of deliberate cheating [*italics added*]. I preferred [*sic*] to run

this risk in order to give feedback. Within the constraints of this method everything was done to discourage cheating or to make it difficult and on no occasion was any cheating detected. Had the results warranted better safeguards these would have been employed after the preliminary experiments. However, it will be seen that elaborate safeguards against subject cheating would have been superfluous” (Blackmore, 1980c, pp. 132—133).

⁵ “Word length confounded with target type” (Blackmore, 1980c, p. 181).

⁶ “The design of the experiment made checking extremely difficult and laborious, so increasing the possibility of errors” (Blackmore, 1980c, p. 181).

⁷ “This experiment was poorly designed in that allowances had to be made for the variation in the number of times each word appeared as target and was chosen by Ss” (Blackmore, 1980c, p. 185).

⁸ “The target pictures were not ideal and could be improved, especially since the relationship between them was unknown” (Blackmore, 1980c, p. 143).

⁹ “In this experiment three key targets and six others were all presented as possible targets to the subjects. This method means that special allowances have to be made for preferences of each type to each target which not only complicates the analysis but may introduce a possible source of error” (Blackmore, 1981b, p. 57).

¹⁰ “Although the subjects were told that the selection of targets was random, they might nonetheless feel constrained to use one of each. . . . This problem of dependence of responses would be much less if more trials were used” (Blackmore, 1981b, p. 57).

¹¹ Inappropriate statistics used (dependence of rankings).

¹² Study labeled “Main Experiment” or “Main Study” in publication, though not differentially distinguished among the dissertation experiments.

¹³ Degrees of freedom and number of subjects are discrepant as “each child took part in each test on a different occasion. A few had a second turn” (Blackmore, 1980d, p. 509).

¹⁴ Experimenter was aware of target pool. Probability value was misreported as .52 (actually .052). Results were said to be qualified due to “only one subject and too few trials” (Blackmore, 1981a, p. 11).

¹⁵ Study reported as “Main Experiment” predates previous “Pilot” study.

¹⁶ Exact date not given in “Schedule of Experiments” (Blackmore, 1980c, p. 135).

¹⁷ Blackmore writes: “It will be noted that in many ways this experiment was less than well controlled. For example it would have been easy for me, as experimenter, to cheat. However, this was only intended as an exploratory study and this was not thought important at this stage” (Blackmore, 1980d, p. 509).

¹⁸ Blackmore served as single subject in this study.

¹⁹ “The results of this exploratory study are included only for the sake of completeness” (Blackmore, 1980c, p. 171).

²⁰ This study was labeled “Main Series” in dissertation (Blackmore, 1980c, p. 172).

²¹ Number of subjects incorrectly stated as 23 in publication (should be 28).

²² Significant result reported in dissertation was omitted from published report.

²³ Reanalysis by Markwick (1988) using proper analysis shows that study retains its significant results.

FOOTNOTES:

¹ The journals searched were *Journal of the American Society for Psychical Research*, *Journal of Parapsychology*, *European Journal of Parapsychology*, *Research Letter*, and *Research in Parapsychology (RIP)*. One experimental report from *RIP* was later published in the *British Journal of Psychology* and is also reviewed herein. No other publications testing the psi hypothesis and meeting the selection criteria were located. [back](#)

² The number 29 is derived from Blackmore’s “Schedule of Experiments” in her dissertation (Blackmore, 1980c, pp. 135—136). This schedule lists each experiment “in its original chronological order” (Blackmore, 1980c, p. 132). No starting dates for any experiment can be found in either the dissertation or subsequent publications. [back](#)

³ Blackmore has written: “I have carried out research into OBEs beginning from the hypothesis that nothing paranormal is involved and the experience is psychological. OBEs have traditionally been part of parapsychology and I believe they should continue to be so regardless of whether any psi is involved” (Blackmore, 1983b, p. 20). [back](#)

⁴ Though the OBE may be a psi-conducive state, as dreams may be, simply inducing the state is not a sufficient condition for psi to occur. Hence, to call OBEs “parapsychological phenomena” may be as inappropriate as calling dreams (or any other altered state of awareness) “parapsychological phenomena.” [back](#)

⁵ Blackmore’s article entitled “The Adventures of a Psi-Inhibitory Experimenter” begins: “I get negative results. Indeed, I have been doing so for ten years” (Blackmore, 1985a, p. 425). I believe this creates the distinct impression that Blackmore is referring to 10 years of experimental work. The jacket of her autobiography (Blackmore, 1986) states, “For more than ten years Susan Blackmore conducted research in ESP, occultism, poltergeists, Tarot cards, and out-of-body experiences.” Blackmore has stated to me in a personal communication (November 12, 1987) that she does not claim to have done 10 years of experiments on psi but 10 years of research on the paranormal. She includes all kinds of research, such as that on OBEs and on checking up on spontaneous cases (such as poltergeists). Thus, the claim of “ten years of research” is a form of “credentials inflation” if we are seeking to consider scientific evidence regarding psi research. [back](#)

⁶ Cognitive dissonance is a social psychological construct that predicts that when faced with contradictions between beliefs, psychological tension will develop and such tension may be relieved by the person changing his or her beliefs. In Blackmore’s case, the contradiction between her choice to invest a large portion of her life to become a doctor of parapsychology is in conflict with the fact that she has been a failure within that discipline (if success is defined by producing research that supports a psi hypothesis). Her response has been to reduce the dissonance by becoming a proponent of a parapsychology without the psi hypothesis (“I’m not a failure, the psi hypothesis is wrong”). I discuss this notion in more detail in my review of her autobiography (Berger, 1988). [back](#)

⁷ Marcello Truzzi (1987) points out that the dictionary defines “skeptic” as one who raises doubts and “is meant to reflect nonbelief rather than disbelief” (p. 8). Thus, the term seems inappropriate to describe Blackmore’s current position. [back](#)

⁸ To aid the reader in identifying the different references derived from the dissertation experiments, a mnemonic word will follow each reference. [back](#)

⁹ This is contradicted by her earlier statement that “three pilot studies were carried out. Because these studies suffered from various flaws they are only described in outline here” (Blackmore, 1981b, “Errors,” pp. 54—55). [back](#)

¹⁰ Experiment 3 in her Table 6 actually refers to “Main Study” [Experiment 4]. [back](#)

¹¹ I have mentioned, in more than one instance, the importance of Blackmore serving as single subject in her own psi experiments. She has publicly stated (see, e.g., Blackmore in Shapin & Coly, 1985, p. 94) that she had never had “an experience of psi.” [back](#)

¹² Though the intent of this analysis was to examine only published reports, her unsuccessful Ganzfeld study is so frequently cited by Blackmore that it was included herein. [back](#)

¹³ The introduction of this journal article states that “five experiments were carried out and are reported here” (Blackmore, 1981a, p. 9), whereas 5 “preliminary” experiments are reported followed by Experiment 6 (reported as “Main Experiment”). [back](#)

¹⁴ The two experiments that are missing from Blackmore (1980a), experiments 8:6 and 8:8, can be found published elsewhere. The former is reported as “Main Experiment” in Blackmore (1981a) and the latter appears as “Main Experiment” in Blackmore (1980d). [back](#)

¹⁵ My count of p values in Chapter 8 yields 3 out of 33 as significant. I assume Blackmore’s 34th value was attached to a correlation where $r = 0$ and no p value was reported. Some of the p values were definitely not independent, for example, where both a t and z score were calculated on the same data and both p values reported (Blackmore, 1980c, p. 212). [back](#)

¹⁶ Exact probability was reported in a Table as .02. [back](#)

¹⁷ In the dissertation version, she wrote that the conclusion was “invalid without further research” (Blackmore, 1980c, p. 144). [back](#)

¹⁸ There are signs of declining scores over time, within-series consistency of scoring (e.g., significant overall ESP hitting in 6 unpublished studies from her dissertation experiments), and significant differences between experiments using Zener cards vs. words as targets. [back](#)

¹⁹ Blackmore states, for example, that the unrepeatability of psi should be taken “as a reason for rejecting the hypothesis of psi. I hope to persuade you [that this is] . . . the only viable solution if we are to have a thriving science of parapsychology in the future” (Blackmore, 1985b, p. 183). [back](#)

²⁰ Blackmore was recently elected a Fellow of CSICOP (Skeptical Inquirer, 13, 1988). [back](#)

[Back to Psi Explorer website](#)