COMMENTS

Second Report on a Case of Experimenter Fraud

By J. B. Rhine

It has been about fifteen months since Dr. Walter J. Levy, Jr.,[1] (W.J.L.) was discovered to be falsifying his test results in a PK experiment with rats at the Institute for Parapsychology, of which he was then Director. He was caught by three fellow staff members, James W. Davis, Jim Kennedy, and Jerry Levin, all of them members of his own research team. As soon as they considered the evidence conclusive, the three men who discovered the deception reported it to me; and after a short but decisive interview with W.J.L., I received his resignation.

Because W.J.L. had published many research papers and his work was widely known here and abroad, it seemed of first importance to warn readers against reliance on any of his research reports. The known evidence against him at the time of the exposé (June 12, 1974) concerned only one of his several lines of research, that in which rats were tested for PK, or psychokinetic ability. The experiment involved the implantation of an electrode in the pleasure center of the animal's brain to provide a means by which random electrical stimulation might make possible the arousal of the animal's desire to increase gratification by influencing (via PK) the rate of stimulation. W.J.L. himself would admit to no other falsification of results, but it seemed advisable to regard as "not acceptable" all reports of his research and all those in which he had played any part (Rhine, 1974b).

In my first report on the W.J.L. affair, I said that there should be a further report on later developments of the case. It is in compliance with that commitment that I now review what has happened

[1] The first report (Rhine, 1974b) used "Dr. W." as the public identification of Dr. Levy in the slender hope of protecting the interests of innocent people in Levy's nonparapsychological circles. The wide general publicity anticipated at that time can be assumed to have passed. The use of Levy's full name is necessary now, of course, in order to identify his work, which has all been disqualified. A complete list of his published reports is given in the Appendix to this paper.

The author is especially indebted to Jim Kennedy, James W. Davis, H. Kanthamani, and Judi Taddonio for their assistance, and as always to the editors.

thus far. First, a statement will be made about what is known of the extent of the dishonesty involved. Second, a roundup will be given of the attempts made to repeat the W.J.L. experiments and how the various projects stand now. Third, I will try to summarize the reactions thus far registered regarding this tragic affair. Fourth, an evaluative word will be offered as to the consequences to the field.

I. The Known Extent of the W.J.L. Falsification

It was a natural question in the minds of all concerned as to how far the cheating exposed in the one experiment with rats extended throughout the rest of W.J.L.'s work and what, if anything, was left unaffected. Consideration was given at once to the possibility of making such a discriminative evaluation. It was reasonable to think that, with all the precautions that had been developed—especially since the elimination of the risk of fraud had long been on our minds—there might be some experiments by W.J.L. in which the conditions ruled out all possibility of dishonesty. It was soon realized, however, that although the precautionary test conditions had been intended to be the best that had been developed in psi research, there was always in actual practice some possibility still open to W.J.L. in his position as Director of the Institute and as a person who thoroughly involved himself in all that was going on. In fact, no one could say that the double-blind conditions, even with a design involving two or more experimenters in each experiment, had been so thoroughly observed that they had completely eliminated all possibility of dishonesty.

It was a unique situation. W.J.L. had been an extremely enterprising and industrious worker. As the chief administrator of the laboratory and an active leader in almost every aspect of its activities, he well knew all the limits and loopholes in a system he was masterminding seven days of the week—often sleeping in the attic while overseeing the 24-hour automatic testing of the animals. In view of these circumstances, there remained none of his work of which it could be said that he himself could not *possibly* have had any opportunity to manipulate the results that had been published. They all had to be put on the same level of uncertainty and unacceptability. (See Appendix.)

I think it can readily be seen that in the concern over an adequate reappraisal of his research there was nothing to do but to write it all off; and this I have done, at first only warning against any reliance on the W.J.L. reports for the time being. But on further reflection it became increasingly clear that the whole of his work was irretrievably lost, even without knowledge of the exact extent of the deception. In any case, such knowledge is not completely obtainable.

However, among individual members of the staff, especially those whose vigilance led to the discovery of the fraud in the first place, there remained unsatisfied questions. These led to some developments indicating that W.J.L.'s deception was much more extensive than had first been discovered. In one of these instances, first encountered by Jerry Levin, some indications turned up as to how W.J.L. could have manufactured results in the gerbil experiments on precognition. He could have done it by shorting one or the other of two wires that would have allowed him to manipulate the random target sequence as he watched the recording of the animal's behavior. It was W.J.L.'s uncalled-for presence at the apparatus with the wires in hand that raised the first question for Levin and led to the exposé. Levin later on observed by accident the presence of fine scratches such as the wires would make if shorted to the aluminum panel.

Another such discovery was made by Jim Kennedy in an examination of the records of W.J.L.'s automated maze tests of clairvoyance with human subjects, the results of which were stored by the computer. This experiment, as devised by W.J.L., had first been conducted with paper-and-pen methods but later two series had been automated with the computer. Kennedy found that W.J.L.'s published results were not at all those recorded for the same experiments on the computer. (The original data were not significant.) Apparently a completely different set of results had been invented.

Still another discovery was made by Kennedy who had earlier done some PK testing under W.J.L.'s supervision using chicken embryos as "subjects." The aim had been to test for possible PK influence exerted on randomly operated heat lamps by the eggs (located as they were in a cool chamber). It had been noticed that there tended to be long strings of hits and an increased trial rate when the scoring was significant. The easiest way to manufacture high scoring in the egg work would produce such effects as these. When the strings were deleted from the data, the scoring rate was at the chance level. Similar stringing was produced as an artifact of W.J.L.'s admitted manipulation of the rat implantation work. Also, the records of egg research which W.J.L. conducted before Kennedy started working with him showed the presence of astronomically significant strings (of about 100 hits where the probability of a hit was ½), yet W.J.L. never mentioned them.

With all the grounds we have for withholding acceptance of W.J.L.'s work, we cannot—and certainly need not—draw a conclusion about the actual extent of the deception in the whole of his five years of reported research. But, as I have already indicated, no salvageable series or section of sound and significant results has thus far been isolated in which W.J.L. was not around somewhere within reach of an unprotected link in the chain of reliable controls. As will be seen further on, the safest and best type of evidence of psi emerges in the form of incidental hidden "signs of psi" (Rhine 1974a) which the experimenter could not himself have anticipated at the time. However, none of these exceptional fraud-proof signs has so far been found in W.J.L.'s work.

II. The Attempts to Replicate

Almost immediately following W.J.L.'s exposé last year, a staff conference was held to consider the desirability of continuing the experiments he had been conducting on which he had reported successful results. The research assistants who had been working with him all seemed to be disposed to see what they could do on their own; there was no immediate readiness on the part of anyone to leap to the sweeping conclusion that everything W.J.L. had had his hand in was fraudulent. It is true, the whole atmosphere of the Institute had considerably changed; but on the whole, and especially among the young people who had worked closely with W.J.L., there was willingness to grant his claims a fair trial.

Difficulties soon arose, however, as the actual situation was fully realized. Jim Kennedy, who was co-experimenter on the rat implantation experiments in which W.J.L. had been caught cheating, was not able to obtain the responses from the rats W.J.L. had reported. On finding that the actual technical basis of the electrode implanting and other details were faulty, the case for W.J.L.'s work seemed increasingly doubtful, and all attempts to pursue this project further were abandoned.

As these comments were first being written I learned from Dr. Helmut Schmidt of Mind Science Foundation, who was here at the time of the W.J.L. exposé, that he has since "spent two months working with rats with implanted electrodes. Six of ten rats showed good self-stimulation if they could induce the stimulus by pressing a lever. Subsequent PK tests under varying conditions . . . did not encourage [him] to continue the tests."

My attention has also been drawn to some possibly related (but unpublished) research by Dr. Gary Heseltine (a recent graduate of

Wayne State University Medical School, Detroit, who is now at Science Unlimited Research Foundation, San Antonio, Texas). This work is intended to be a test of psi exchange between two rats. In Dr. Heseltine's words, "preliminary work indicates that the EEG of one animal can be altered when the occurrence of certain EEG events results in [pleasurable] electrical stimulation of the brain of a second animal." Heseltine intends to continue the research.

Problems were also encountered by Kennedy and by Douglas Richards, who together took charge of the tests of PK ability in young chicks and fertilized eggs after Levy left. Although both experimenters had had successful results while W.J.L. was overseeing the research, neither could replicate their previous work after he had gone. While Kennedy was away, Richards, working alone, carried on the work into the winter, when hatching difficulties were encountered. With Richards on part time (while in graduate school) and with poor hatching rates and chance results, the project was eventually abandoned.

However, I learned recently that biologist John Randall (Leamington College, England) had continued his PK tests with this procedure which he had begun while W.J.L. was still here. Randall informs me that he has now completed six series of tests on which he is able to make an interim report. The investigation has fairly consistently given slightly positive results with a marginally significant total, and he plans to complete another six series.

It should be recalled here that W.J.L. did not originate the research on animal PK. That innovation had been first reported by Dr. Helmut Schmidt while he was at the Institute, mainly on the basis of his experiments with cockroaches (Schmidt, 1970). Schmidt, however, had abandoned the effort to develop a satisfactorily repeatable animal PK experiment. He had also given the eggs a try while W.J.L. was still present at the Institute (coming and going from medical school); but he had generally failed to replicate at W.J.L.'s level of success. These exploratory attempts, as well as others made with some success with still other species of animals, did not reach the stage of meriting publication. Some of this animal work, however—especially the Schmidt tests and now the Randall interim report—will help to keep the question of animal PK in the focus of interest regardless of the work of W.J.L.

Two other published lines of research by W.J.L. have also been repeated to some extent. One of these was the first psi experiment taken up by him (and three co-workers) as an attempt to replicate the experiment with small rodents in tests of precognition of the type originally reported by Duval and Montredon (1968). In fact, most of W.J.L.'s claim to recognition in parapsychology was based on his five years of almost unbroken success (as reported) with this line of research. Especially outstanding in these precognition tests with small rodents was the long succession of significant results in the random behavior trials and in some further behavioral categories which W.J.L. himself developed. As we surveyed these many replications in the gerbil experiments and some closely related findings with rodents already reported by others (Parker, 1974; Schouten, 1972), there seemed to be some prospect of successful replication at the Institute.

The man best prepared to continue was Jerry Levin, who had worked with W.J.L. in the rodent tests, although not actually on tests of precognition. He had not, however, yet conducted a psi experiment of his own. The important point was that he was here, was willing, and that he already knew something of how W.J.L. had tested the animals for precognition. Over the months that followed, Levin conducted five series of tests with gerbils, using both the wheel and the box apparatus; however, only one of these five series gave significant results, although there were suggestive effects that would help to justify continued interest. But Levin (1975) concluded that, with the time and effort expended, he was not able to confirm W.J.L.'s findings.

Shortly after Levin had finished his five series and had given up the effort to replicate W.J.L. on gerbil precognition, a former assistant of W.J.L., James Terry, then located at Maimonides Hospital in Brooklyn, expressed an interest in repeating Levin's gerbil precognition tests, and he was invited to the Institute to do so. He had earlier joined in an exploratory series of tests of ESP in rats at Mt. Holyoke College with Susan Harris (Terry & Harris, 1975) and had also conducted psi research with human subjects (Terry, 1975a; Honorton & Terry, 1975). During the summer of 1975 Terry carried out four series of tests with gerbils at the FRNM, using the W.J.L. box apparatus. He, too, obtained only one significant series among the four. Again, although it has suggestive value, this work by Terry (1975b) does not constitute independently confirmatory evidence of psi.

However, a number of experimenters outside of the FRNM have presented reports of precognition tests with rodents. W.J.L., as already stated, had been attempting to replicate the original work of Duval and Montredon. After he began reporting his results, other experimenters took up the project, changing only the method of motivating the animals to that of more positive treatment (water and food rewards instead of electric stimulation). Several significant experiments have been reported (Schouten, 1972; Parker, 1974) along with some failures too. Dr. H. J. Eysenck (1975) has recently reported a significant series with a nonaversive level of electrical stimulation. Also, at the FRNM, James Davis has conducted a suggestive exploratory series of precognition tests with rats based on food rewards for hits. This procedure more closely resembles the rodent tests by Parker and Schouten than those by W.J.L. or Duval and Montredon. Davis, after one successful series followed by one at the chance level, is still continuing the project. As usual, an unknown number of failures to obtain evidence may have occurred in other independent attempts to repeat these tests. But since no one can reliably know what may have been done wrong and therefore why they were unsuccessful, they neither prove nor disprove anything of importance at this stage (Rhine, 1975).

The other line of W.J.L.'s research in which independent confirmation has been attempted is the maze test of ESP, which, incidentally, was the only psi research he conducted using human subjects. A repetition of the automated form of maze test has been carried out by James Davis, with results that justify further work; however this has not yet been completed. Also, Jim Kennedy and Kathryn Parker have carried out experiments with the automated maze technique but with a somewhat different procedure; their report, too, is awaiting some clarification of the statistical method before publication.

Still another, and more independent, effort to repeat the maze test was begun before W.J.L. left the Institute. This was conducted by Stephen H. Glidden, Jr., (1974), a student at Suffolk University. In this research, a slight modification of W.J.L.'s paper-and-pen test procedure gave a marginally significant total (p = .025, one-tailed) for the random behavior trials. Other independent research on the maze procedure has been going on elsewhere (again with minor modifications) but has not yet been concluded. Thus the maze technique with human subjects and the random behavior analysis are still on trial, with possibilities that the confirmation can be further improved. W.J.L. had devised it at my request in order to adapt the random behavior technique as a way of eliminating wasted effort in tests with human subjects, as it seemed to be doing in the animal research. There is a fair prospect of continuance of this line.

Where do the efforts at repeating W.J.L.'s projects leave these various research problems? As we see now, nothing definitely confirmatory has so far come out of all the attempts made here since his departure. The results from the Institute are lacking in acceptable significance; and without evidence of psi, nothing reliable and meaningful can be said with reference to W.J.L.'s reports. But while these attempts at replication give W.J.L. little support, on the other hand they have not disproved anything he reported. (It is the evidence of data falsification that disqualifies his work.) The point is that, as far as these failures to replicate go, W.J.L. *could* have done his tests reliably and every one of those who were repeating them *could* still have failed to provide the necessary conditions well enough to produce evidence of psi. This fact regarding attempts to replicate a psi experiment is elementary in scientific parapsychology—it has been well known for decades.

At the same time, it is interesting to note that the more successful attempts at the repetition of W.J.L.'s projects have come from other laboratories. One might almost think that psi-testing ability at the Institute had blanked out as a result of the shock of this exposé; however, that would not be correct. For what it is worth, I will add that most of the currently productive research here has been along lines other than those pursued by W.J.L.

It therefore seemed advisable, after some months of effort at "bailing out" the W.J.L. projects, to abandon that objective as an express commitment. Given a similar situation again, I would not even advise consideration of any obligation to do so in the first place. Naturally the problems on which W.J.L. worked still remain as important as ever, although deriving no support from any of his reports. In time they may all be taken up again by others and probably even here at the Institute. Several of them, as I have indicated, are now being actively pursued elsewhere.

However, it is interesting to note that in all of these *independent* repetitions in other centers some modification of procedure was made. These changes could be quite important to the experimenter's motivation in giving him a special personal affiliation with the experiment. Those working here in W.J.L.'s own former laboratory, on the other hand, may have lacked this motivation for the undertaking, especially under the depressing circumstances that prevailed. There was not the freedom to "make it over in their own way" that might have made a difference. (Incidentally, the one W.J.L. project still being pursued here, the maze test, has undergone important adaptations too.) Naturally these are speculations. In a science as non-exact as parapsychology still is, we cannot yet expect to infer reliably from one experiment to another just why one was more successful than another in producing—or in failing to produce—evidence of psi.

As it is, repeating W.J.L. closely and getting nonsignificant results told us really very little. We already knew that psi experiments are not yet easily repeatable; it is too easy to do them wrong in too many ways. But we are at least reminded that we have not reached the stage of successful control over psi that W.J.L. was leading us to expect. It is encouraging, however, to see that whatever our own setbacks here, Europe (in the United Kingdom, Holland, and France) has been taking the lead in psi testing with animal subjects. The goal of reaching the guinea-pig stage in psi research is still a worthy objective.

III. Reaction to the Exposé

In general, the way parapsychologists received the news of W.J.L.'s dishonesty was, like our own, one of shocked surprise and amazement. *Why* did he do *that*? How could he even have felt the need to do such a thing after all the success he had had? What a careless way to cheat—not even clever! Was he in his right mind? Had he been overworking? Was he perhaps in poor health? Had he given other signs of unreliability or instability? Would this not ruin his career? How could he possibly stay in parapsychology? Etc., etc.

Never before had any such sympathetic response been received at our research center. Our colleagues in the field—that is, those who communicated, as many did—were generous and understanding; they were ready to help if help were needed. There was strong approval of the candor and dispatch with which the news was released and the warning given to those who especially needed the information. I have appreciated also the way people outside the field responded to the news. This fair-minded attitude remains one of the strongest impressions I retain of the after-effects of the W.J.L. case.

The reaction toward W.J.L. himself was, in proportion to the age of those who commented, also more one of sorrow than of anger; and the hope was expressed by many that, with all his ability, he could somehow salvage an effective and useful career out of the ruins of this one; but everyone considered it would have to be in another, more objective field than parapsychology. (And now it is.)

It was gratifying, however, that no one in or out of the field ever even asked the gloomy question: "Is this going to ruin parapsychology?" A few did inquire if it would be likely to turn many people away from the field; I could not answer with certainty, of course, but I did not think it would. The forthright exposé seemed rather to increase public confidence. To the question whether it would discourage donors from helping to finance research, happily I could say that one of our leading sustainers, Dr. Marie Higbee, promptly made a special and very generous gift, just because she thought our morale would need the support at that particular time.

Finally, I want to register a word on my own personal reaction when, to face this crisis on June 12 last year, I was abruptly called from the relative retirement into which W.J.L.'s administration had finally enabled me to escape. I was inevitably required to assume again a share in the many burdens left by his sudden withdrawal. At the time and stage of life I had reached, it was a most unwelcome adjustment—one from which I have been hoping ever since to escape at the earliest honorable moment. Thanks to a devoted and capable staff, the FRNM had been able to survive the upset and sturdily display the signs of sure recovery.

Yet the tragedy of this crisis in parapsychology will not let me lose the resolve to try to see that an ugly development like this one shall never again strike our field. One can surely hope that the kind of drastic effort all of us in parapsychology must now undertake may be all the more a united one just because W.J.L.'s tragic blunder has brought the issue home to us all as nothing else conceivably could have done. It should now be possible to obtain a more concerted effort all around than could otherwise ever have been enlisted. A special obligation to this task I must myself accept since obviously, among all who shared some responsibility for W.J.L.'s meteoric career in psi research, I, more than anyone else, could have used greater vigilance and wisdom by pursuing a policy which would not have permitted a climax so destructive both to the science and the man himself.

IV. Consequences to Parapsychology of the W.J.L. Affair

The long-term effects of the W.J.L. case on parapsychology will probably depend mainly on the unity of response among colleagues in the field. It should therefore help in the long run in this unification if the fraud problem in general is kept in the focus of attention in the laboratories and the journals. This is not likely, I think, to be overdone and to create a morbid distraction to normal research. It will probably not arouse that much interest.

However, the fraud issue, even in its most generalized form, is but one of several major problems in keeping psi testing on a high level of security. The multiple aspect of research reliability was outlined at the beginning of an article on "Security versus Deception" (Rhine, 1974a). At the moment, however, the primary topic is that of fraud, and the specific issue is that of experimenter dishonesty. Yet, even now, it is still not just the deception of the W.J.L. type that we need to consider here. The very timing of this affair at the FRNM ties it up with the broader problem of possible experimenter dishonesty beyond this individual case and further complicates the search for adequate solutions. In fact, we can best regard the W.J.L. case as one that pointedly draws attention to a larger insecurity situation which, as I have already indicated (Rhine, 1975), was emerging well before the W.J.L. exposé took place.

Whatever the consequences of this fact are to be and whatever we are to do about them will likely be better understood if we take something of the earlier background into account. When we do so, it is much more understandable that we find ourselves in our present state of special concern over fraud in parapsychology. It is well known that Henry Sidgwick, the first President of the S.P.R., warned that when critics found no other points open to attack in psychical research, they would charge the investigators with fraud. That this would be the final recourse was, of course, logical. I can recall that in the 1930's experimenter fraud was something we had to force ourselves to take seriously. It was only after we had arrived at a relatively conclusive stage of the ESP tests that McDougall advised me to introduce the two-experimenter test design to meet the anticipated charge of fraud. The experimenter deception which this precaution was designed to exclude was not nearly so immediately urgent as had been the exclusion of sensory cues, the proper mathematical evaluations, the safeguarding of records, and the interpretations of the results. Even the critics did not begin to hammer effectively on the experimenter-fraud question until George Price's article appeared in *Science* (1955), and Price himself argued in effect that since no other criticism was crucial, the conclusion had to be fraud. So this was the predicted last resort of the skeptic.

I am reviewing this familiar situation as background for a closely related point that needs attention here: it is that the very same development that finally closed in on the critics' options, cornered the experimenter himself. At earlier stages of method, the experimenter could alternatively obtain spurious results under test conditions that allowed several uncontrolled factors, but one by one these gaps in method had been closed.

Those who understand science can see that this has been a progressive development in parapsychology in which not only is the unyielding critic driven by elimination into a last-ditch alternative of fraud, but also the experimenter who is unsuccessful in obtaining evidence of psi in his tests is left with the single choice of accepting his failure and trying again or of falsifying his results.

I must interrupt at this point to comment that a few colleagues insist we are *making it too hard* for the experimenter by this gradual improvement of our methods so that at the end he has no alternative to the final evidence we are seeking. They hint that we are thus helping to make him cheat by our very measurement of success. Several spokesmen of this viewpoint are proposing that we lower the standards and abandon our reliance on the measurements of significance. They forget that in present psi testing this requirement of significance is the only way we *can* measure success. Might we not just as well say, "Let us freely allow sensory leakage"; or "Let us use no safeguards against recording errors." Surely we need not be frightened off our course of methodological advance by the fraud issue any more than we have been at other stages of the development of psi methodology. The way to help the less resolute experimenter is not to weaken the safeguards that make conclusions possible, but to strengthen the experimental program against error in every way.

My special point in drawing attention to the fact that dishonesty was the *only* option W.J.L. had was that, with the psi test methods what they are today, this alternative of deception was more easily and conclusively identified as such when the real evidence he had expected failed him. It was not hidden in a loose methodology that would have left the matter inconclusive and confusing. Bad as the situation was, therefore, we can still appreciate the advance in methods in parapsychology that facilitated the exposé; we have therefore reached the crossroads which Price, Hansel, and others have long since recognized—the point where, from the view of extreme skepticism, "it had to be fraud." Very well; if that is where we are, it is to our advantage that now we have a clearly drawn contest, and I think we are better prepared to face it than ever.

But again, it is not just the critics' last stand; it is not only the experimenters' either. It involves the whole field quite as much. In the past the trouble has been that no more attention was paid to the question of experimenter fraud than the current urgency of a given case demanded. The principal lesson of the W.J.L. incident is to remind us of the lukewarm concern that has been felt generally throughout our field over the question of experimenter dishonesty.

Now we must be ready, more than ever before, to give the attention this issue requires. It is mainly now a matter of how much we can accomplish toward that end while the urgency is still widely and keenly felt.

It is true, much should be accomplished in time by the kind of increasing safeguards that I have been discussing in recent years, and especially as others join the endeavor. Recognizedly, these past advances in methods, even while they do make fraud the only alternative to successful psi demonstration in a well-designed experiment, will also make it equally more difficult for an experimenter to take that option and falsify his results. Better still, these advances in safeguarding methods should make the temptation to cheat less appealing and still more unrewarding as psi-testing designs improve. But I am coming to question, as I ponder the reactions to the W.J.L. incident, whether we can afford to wait for this rate of advance. I find it easier to ask: "Why *should* we wait at all in this 'last-ditch' stage of indecisiveness? Is it really necessary?"

I have heard it stated at times in the last year that the suspicion of experimenter fraud could be expected to take an indefinitely long time to vanish completely from psi research. Today I do not think it has to be that way at all if we really take the problem seriously and all come to know the literature of the field adequately. In fact, I even think it should be possible to bring parapsychology through this final challenge to its security, not only with success but with reasonable dispatch. In making this proposal, however, I am assuming we will still go resolutely on with a vigorous development of the methods we already know, giving full attention to the safeguards we now have at our disposal for reducing the risk of experimenter deception. But it is time for a more conclusive step.

A Roundup of Fraud-proof Evidence is Essential

Since circumstances have made a mountain of apprehension over the fraud issue, an equally mountainous formation is called for to represent all the many types of fraud-proof evidence the field of parapsychology has to offer, most of which are little known as such even within the field. It is fair to say, too, that much more may yet be discovered that is still hidden from us all. This will call for a determined search; first, among the stores of existing records available for reexamination; and second, in new types of researches which lend themselves well to the purpose, as (judging by the past) many may be expected to do. It will suffice now, however, merely to indicate some of the principal kinds of evidence already known, to which the term "fraud-proof evidence of psi" can be applied. The category itself and its qualifications and limits can be refined later. The following few types will, I think, serve the purpose of introducing the wealth of this kind of material available without delay.

First, I would propose an examination of all available reports by scientific authors who for some reason wished not to publish their work, but who conducted acceptable experiments for their own satisfaction. As an example, I will cite the case of a report by an eminent scientist who was at the time the head of an internationally known institution. Many years ago he showed me his paper but felt he had to withhold it from publication in order to avoid embarrassment to his colleagues. He was, however, persuaded to publish the report of the experiment if all marks of identification of the author were removed; and this was done. The results were significant and were even repeated later.

This type of incident has other variations. Sometimes the fear of losing his own status has kept an author from publishing, although he was willing to circulate a report privately. Or again, the author may earlier have become known as a critic of psi research and was reluctant to admit that he had later conducted an experiment and that significant results had been obtained. Still other varieties of this kind of situation, one in which it would not be reasonable to suspect the fraud motive, can be assembled. In many cases the actual names could now safely be used. The point these situations have in common is that it would not make sense even to suggest that there might have been any will or intention to fabricate the data. To be sure, such researches have to be of good quality in other respects as well.

Second, I would draw attention to another type of incident that has occurred from time to time in the editorial offices of the *Journal of Parapsychology* or in the correspondence of the Institute for Parapsychology. In a typical case, a report of a psi experiment is received for publication from an author who is rather new to parapsychology. The results may or may not be reported as significant. In the course of editorial consideration of the paper it becomes evident that the results lend themselves to a type of analysis which is unknown to the author but which may be independent of the analysis he has already made. Or the author's attention may be drawn to a finding reported by another experimenter which was then found by the editorial staff to be confirmed by the data of this newly submitted report. When such an application is clearly new to the author and is based upon an earlier finding by someone else, it can not only be given full acceptance, but it is all the better for having been brought to the author's attention by an independent analyst. The author can, of course, then make an independent check on the analysis from his own records, i.e., to confirm the editor's findings. The fraud-proof value of these cases covers a considerable range of security, often well beyond that of experimenter deception; some are of the ideal type in which one effect is found to be a completely independent replication of an experiment made under a very safe set of conditions, as, for example, when independent analyses are made on duplicate records.

Third, even better than the preceding types, although somewhat more rare, are examples in which already published work may be reexamined long years after publication for a significant distribution of hit patterns in the data, free of any conceivable motivation on the part of either subjects or experimenters. Take, for example, the familiar telepathy tests by G. H. Estabrooks at Harvard in 1925 (Estabrooks, 1961). It will be recalled that he obtained positive results in card guessing in three out of four series of tests. In the fourth, in which the conditions were changed to a greater distance between subject and target, the group of subjects scored well *below* mean chance expectation. Moreover, in all four series there was a marked top–bottom decline of scoring rate in the 20-trial runs. But Estabrooks himself had been aiming at high-scoring total results, and these were the sole basis of his conclusion. He made no evaluation of the significance of the other features although he noted them; there was no precedent then known to him for doing so. Many years later, when I looked into his data for the type of results we had been getting in my lab at Duke in somewhat similar ESP tests (e.g., psi-missing and declines), I found the negative deviation of hits in his fourth series to be significant. Also, the decline in the runs throughout all four of his series was quite definitely significant in the top–bottom difference in scoring rate. This had become a standard way of evaluating decline effects, but that was long after the date of the experiment. It seems quite reasonable to think that neither Estabrooks nor his subjects could have been motivated to fabricate these peculiar results; rather they were safely attributable to the ESP process itself and in time came to be recognized as typical signs of psi.

Other variations, too, have been found that depend on peculiarities of the psi process overlooked by the original experimenter. One of the most complex examples of this third type is that of the PK record book of M.P.R., a Guilford College student of psychology (Rhine & Reeves, 1943). This experimenter's original tests of dice-throwing were self-tests, not intended for publication;

but because they were faithfully recorded in a very systematic way and were strictly routine, they were reexamined years after the experiment was over when a search was being made for position effects at the Parapsychology Laboratory at Duke. A hit-distribution analysis conducted by Betty Humphrey was applied to six breakdowns of the data in which decline effects could be expected on the basis of other, previously analyzed work. Of these six breakdowns, five showed the significant differences expected (for example, the top–bottom difference of the hits in the run). The consistency of the patterning was unquestionably a remarkably evidential effect which, under the circumstances, ruled out any conceivable question of dishonesty in the original tests. An independent analysis followed Humphrey's, and others could still be repeated. (The nature of the records was such that trickery in the analysis would almost surely have been discovered. For example, erasures and other alterations are detectable by the experienced analyst, and the records were in a bound composition book.)

Fourth, by far the most imposing body of evidence of the fraud-proof type also occurred in the PK branch of the field. By 1944–1945 we had accumulated a large number of research reports from our own laboratories and were examining them for methods of appraisal that would get beyond the great variety of conditions under which they were conducted. The most common feature among all 24 of these reports was the record sheet, a more-or-less standard form on which the results of dice throwing by one means or another could be recorded for the most part in regularly structured rows or columns on the page. In many experiments the page was divided into four, six, or eight subunits or sets. Eighteen of these reports had pages that were uniform enough to be subdivided into quarters, sometimes omitting median runs, columns, or sets to get uniformity. Twelve of the 18 had these regular sets on the record page.

This afforded a substantial common base of hit distributions with enough uniformity for generalization, and the expected declines in the column or row from the first to the last trial could be measured both ways by the quartering of the page (and the set). The quarter distribution (QD) of the page then became the basis of appraising hit distributions in the 18 series, and an *independent QD analysis* could be applied to the sets in 12 out of the 18. The test of significance to be applied was the difference between the upper left and the lower right quarters, both in the page QD's and the set QD's. These diagonal declines, as they were called, proved to be highly significant and remarkably consistent. Sixteen of the 18 series showed the decline and the *declines were even more significant in the set* than on the page. The point here, of course, is that none of these original experimenters back in the earlier years when the tests were conducted had any idea such an analysis would ever be made. Most of the series had shown significant total scoring above the chance mean. About those results, in some cases more than others, we could have raised questions of accuracy of recording, of bias in the dice, and of unreliability in the experimenters. But after these QD analyses were made, no human weakness or bias in the apparatus could be found that could have produced these declines. Again, Dr. Humphrey did most of the analyses; but prior to issuing an invitation to qualified outside observers to repeat them, I invited Dr. J. G. Pratt to the Laboratory to make a complete independent analysis. Pratt's analysis confirmed Humphrey's very closely.

At the Duke lab we recognized the fraud-proof conclusiveness of the QD analyses; but in the decade that passed before the fraud issue was first raised in a big way by G. Price (1955), the passage of time had allowed the QD's to be largely forgotten. They were not taken seriously by the critics. Yet, meanwhile, they have lost nothing of their antifraud potential.

These four kinds of special evidence against psi fraud do not exhaust the list of possibilities by any means; but they will, I think, suffice to show that this approach to a final settlement of the establishment of psi occurrence has a conclusiveness of its own. We need not leave the contest to the experimental attempts at fraud control where possibilities of the experimenter's motivation to deceive will also have to be dealt with. I am reasonably sure these fraud-proof types of evidence, if now given precedence for a showdown on the establishment of psi, will fill the need for the scientific mind that considers it. But it will take a thorough effort to draw that issue forcefully. Again, however, it must be repeated, this is not to abandon the effort to build up the control system that will eliminate experimenter deception along standard experimental lines. The evidence of hidden signs of psi, while more conclusive when it applies, is less adaptable (at least as yet) to all test situations. So for the present, we need both lines of control against fraud.

Once the case for psi is conclusive on the basis of the hidden type of evidence, there should be changes welcome to all. Beyond doubt the revolutionary character of psi makes decisive proof exceedingly hard to produce, in contrast to that in the physical sciences. Since there is already some evidence that the subject's belief in psi occurrence affects his test performance, this should be important to the scoring levels. Perhaps experimenters, too, would be affected by a generally relaxed acceptance of the fact of psi ability. We might anticipate that in time there would be little more need for concern about honesty than in any of the other psychological branches of testing and research. Most helpful, perhaps, would be the reduction of skepticism and indifference on the part of those who should be looking into the bearing of psi on their own fields. The future of psi research depends more on this sharing of interest from other disciplines and sciences than upon anything else.

The whole program of parapsychology should be expected to respond to the confidence springing from the assurance that the claim of psi has passed all its tests and that there are no last details which can provide the scientific mind with an excuse even for suspended judgment. More confident steps should then be possible to provide parapsychological education of the highest quality, to obtain adequate economic support, and even to consider practical application over a wide range of possibilities. All this may come in time; but the timing itself may endanger the outcome if left to the unguided circumstances of evolution. So I would say that, with a full appreciation of the major impact of the W.J.L. exposé, it would be best not to forget it until it has spurred us all to clear the terminal barrier to the goal of ultimate establishment.

Appendix

Work Authored By W. J. Levy, Jr. 1969-1974

1969

Levy, W. J. Motor skill and psychokinesis. *Journal of Parapsychology,* 1969, **33**, 325. (Abstract)

1970

Levy, W. J. Effect of the test situation on precognition in mice. *Journal of Parapsychology,* 1970, **34**, 278. (Abstract)

Levy, W. J., & André, E. Possible PK by young chickens to obtain warmth. *Proceedings of the Parapsychological Association,* 1970, No. 7, 8-9. (Abstract). Also in the *Journal of Parapsychology,* 1970, **34,** 278-279; 303 (Abstracts)

Levy, W. J., & McRae, A. Precognition in mice and gerbils. *Proceedings of the Parapsychological Association,* 1970, No. 7, 9-10. (Abstract). Also in the *Journal of Parapsychology,* 1970, **34**, 279; 303-304. (Abstracts)

1971

Levy, W. J. Possible PK by chicken embryos to obtain warmth. *Proceedings of the Parapsychological Association,* 1971, No. 8, 25-27. (Abstract). Also in the *Journal of Parapsychology,* 1971, **35**, 61; 321-322. (Abstracts)

LEVY, W. J.; MAYO, L. A.; ANDRÉ, E.; & MCRAE, A. Repetition of the French precognition experiments with mice. *Journal of Parapsychology*, 1971, **35**, 1-17.

LEVY, W. J., & MCRAE, A. Precognition in mice and jirds. *Journal of Parapsychology*, 1971, **35**, 120-131.

1972

LEVY, W. J. The effect of the test situation on precognition in mice and jirds: a confirmation study. *Journal of Parapsychology*, 1972, **36**, 46-56. Also in the *Journal of Parapsychology*, 1971, **35**, 60-61. (Abstract)

1973

LEVY, W. J. Random behavior in an experiment with humans. *Journal of Parapsychology*, 1973, **37**, 1-12.

LEVY, W. J.; ARTLEY, B.; WILLIAMS, C.; & OWENS, B. Effects of factors interpretable as high and low stress states on precognition in small rodents. *Research in Parapsychology, 1972.* Metuchen, N. J.: Scarecrow Press, 1973, Pp. 159-162. (Abstract)

LEVY, W. J., & DAVIS, J. W. A comparison of variable- and fixed-trial intervals in the rodent precognition work. *Journal of Parapsychology*, 1973, **37**, 201-209.

LEVY, W. J., & DAVIS, J. W. Introduction of an activity-wheel testing cage into the rodent precognition work. *Journal of Parapsychology*, 1973, **37**, 253-277.

LEVY, W. J., DAVIS, J. W., & MAYO, L. A. An improved method in a precognition test with jirds. *Journal of Parapsychology*, 1973, **37**, 83-96.

LEVY, W. J., DAVIS, J. W., & TERRY, J. C. Two possible sources of ESP information in the rodent precognition work. *Journal of Parapsychology*, 1973, **37**, 189-200.

LEVY, W. J., & TERRY, J. C. Further study of the wheel testing cage in the rodent precognition work. *Journal of Parapsychology*, 1973, **37**, 323-333.

LEVY, W. J., TERRY, J. C., & DAVIS, J. W. A precognition test with hamsters. *Journal of Parapsychology*, 1973, **37**, 97-104.

1974

LEVY, W. J. An automated maze test with random behavior trials by humans. *Journal of Parapsychology*, 1974, **38**, 27-46.

LEVY, W. J. Possible PK by rats to receive pleasurable brain stimulation. *Research in Parapsychology, 1973.* Metuchen, N. J.: Scarecrow Press, 1974. Pp. 78-81. (Abstract)

LEVY, W. J.; ARTLEY, B.; MAYO, L. A.; & WILLIAMS, C. The use of an activity wheel based testing cage in small rodent precognition work. *Research in Parapsychology, 1973.* Metuchen, N. J.: Scarecrow Press, 1974. Pp. 71-74. (Abstract)

LEVY, W. J., & DAVIS, J. W. A potential animal model for parapsychological interactions between organisms. *Research in Parapsychology, 1973.* Metuchen, N. J.: Scarecrow Press, 1974. Pp. 28-31. (Abstract)

References

Duval, P., & Montredon, E. ESP experiments with mice. *Journal of Parapsychology,* 1968, **32**, 153-166.

Estabrooks, G. H. A contribution to experimental telepathy. *Journal of Parapsychology,* 1961, **25**, 190-213.

Eysenck, H. J. Precognition in rats. *Journal of Parapsychology,* 1975, **39**, 222-227.

Glidden, S. H. A random-behavior maze test for humans. *Journal of Parapsychology,* 1974, **38**, 324-331.

Honorton, C., & Terry, J. C. Psi-mediated imagery and ideation in the ganzfeld: a confirmation study. *Research in Parapsychology, 1974.* Metuchen, N. J.: Scarecrow Press, 1975. Pp. 93-96. (Abstract)

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

Parker, A. ESP in gerbils using positive reinforcement. *Journal of Parapsychology,* 1974, **38**, 301-311.

Price, G. Science and the supernatural. *Science,* August 26, 1955, p. 359.

Rhine, J. B. Security versus deception in parapsychology. *Journal of Parapsychology,* 1974, **38**, 99-121. (a)

Rhine, J. B. A new case of experimenter unreliability. *Journal of Parapsychology.* 1974. **38**, 215-225. (b)

Rhine, J. B. Psi methods reexamined. *Journal of Parapsychology,* 1975, **39**, 38-58.

Rhine, J. B., & Reeves, M. P. The PK effect: II. A study in declines. *Journal of Parapsychology,* 1943, **7**, 76-93.

Schmidt, H. PK experiments with animals as subjects. *Journal of Parapsychology,* 1970, **34**, 255-261.

Schouten, S. Psi in mice: positive reinforcement. *Journal of Parapsychology,* 1972, **36**, 261-282.

Terry, J.C. A multiple session ganzfeld study. *Research in Parapsychology, 1974.* Metuchen, N. J.: Scarecrow Press, 1975. Pp. 48-49. (Abstract) (a)

Terry, J. C. A continuation of the rodent precognition experiments. *Journal of Parapsychology,* 1975, **39,** 366. (Abstract) (b)

Terry, J. C., & Harris, S. A. Precognition in water-deprived rats. *Research in Parapsychology, 1974.* Metuchen, N. J.: Scarecrow Press, 1975. P. 81. (Abstract)

Institute for Parapsychology
College Station
Durham, N. C. 27708