Journal of Scientific Exploration, Vol. 13, No. 4, pp. 561–577, 1999 0892-3310/99

Experimental Systems in Mind-Matter Research

ROBERT. L. MORRIS

Department of Psychology, University of Edinburgh
Edinburgh EH8 9JZ, United Kingdom

and

Institut für Grenzgebiete der Psychologie und Psychohygiene
79098 Freiburg, Germany

Abstract — Research into mind-matter interactions (MMI) investigates the possibility of some sort of relationship between the consciousness of living beings and external material systems, mediated by mechanisms beyond those we presently understand. This paper presents a conceptual basis for experimental strategies in the investigation of such effects. Twelve different classes of interpretations of evidence are presented and described. An observer-centered communication model is developed for use in designing MMI experiments and its application described. It is linked to three kinds of ESP designs (restricted choice, free response and somatic) and three kinds of PK designs (discrete outcome, stable system and static system). Some measurement issues in MMI research are raised, relating to evaluation of the null hypothesis and the concept of psi-missing.

Keywords: statistical evidence — *p*-values — meta-analysis — repeatability

Introduction

Research into mind-matter interrelations (MMI), in the sense presented by Atmanspacher in his introduction to this issue, investigates the possibility of some sort of relationship between the consciousness of living beings and external material systems, mediated by mechanisms beyond those we presently understand. Such a relationship is suggested by instances in which conscious, experienced internal (mental) events are inferred by observers to correlate with external, material events, under circumstances which should preclude conventional means of linkage or communication between them. For present purposes, these will be regarded as MMI effects. These instances include spontaneous, unplanned experiences; informal but deliberate attempts to induce such effects; as well as attempts to elicit the effects in controlled laboratory experiments. It should be noted that in some contexts MMI effects suggest causal interaction, thus going beyond the issues raised, for instance, by mind/brain correlations, which may or may not involve a direct interaction.

For such effects to be added to the extant corpus of scientific knowledge, they must be amenable to systematic observation and conceptualized in a way enabling experimentally testable predictions. To conduct experimental studies of MMI, one must demonstrate evidence for the effect, show that it is not readily interpretable through known mechanisms (thus representing anomaly), and produce measurements sufficiently valid and reliable to allow process oriented research to take place. This in turn is necessary to allow the development of models of what such phenomena are, rather than merely what they are not. The rest of this paper will explore a set of existing strategies, drawing from the literature of experimental parapsychology. It will regard experiments as systems, composed of interactive components, and will use the definitions employed by the Koestler Parapsychology Unit of the University of Edinburgh to illustrate one kind of systematic approach.

A Conceptual Basis for Experimental Strategies

When the Koestler Chair of Parapsychology was set up at Edinburgh, its declared intention was to further objective scientific research into: "the capacity attributed to certain individuals to interact with their environments by means other than the recognized sensory and motor channels." We have come to define parapsychology as "the study of apparent new means of communication, or exchange of influence, between organisms and environment" (*e.g.* Morris, 1996). By "new" we mean in the sense of not presently recognized by consensus science. Two main categories have generally been employed in the literature. In extrasensory perception or ESP, the organism appears to receive information or influence from some aspect of the environment. In "psychokinesis" or PK, the organism appears to exert an influence upon some aspect of the environment. ESP is sometimes further divided into telepathy, in which one individual appears to receive information or influence from the thoughts or experiences of another; clairvoyance, in which an individual appears to receive information or influence from a physical object or event; and precognition, in which an individual appears to receive information or influence from a future event or experience. ESP is sometimes referred to as anomalous input, anomalous cognition or receptive psi; PK comparably can be regarded as anomalous output, anomalous perturbation, or emissive psi. Although these two categories can be conceptually distinguished as separate potential capabilities, in practice, as will be seen below, they can be difficult to separate experimentally. Together they are often referred to as psi, with the possible exception of telepathy, should telepathy be construed as an interaction between consciousnesses without matter as an intermediary.

For present purposes, psi effects can be regarded as the same as MMI effects. If we are to draw inferences about new means of interaction, it can be important to understand the various classes of interpretations that can be placed upon instances which appear to represent MMI effects. In each case we are talking about situations in which a mental event is observed to resemble an external environmental event, such that they appear to be linked in some way, and yet there do not appear to observers to be any obvious known means of linkage between the two.

Several kinds of interpretation can be offered for the various observations suggestive of MMI. Below are some of the main general categories of interpretation:

- *Coincidence*. Two events will often resemble each other just by chance. We are notoriously poor judges of how often correspondences in the world around us are likely to occur without any causal connection. Diaconis and Mosteller (1987) give a clear description of the many problems involved in evaluating the role of coincidence in daily incidental events. In controlled studies it is important to define the domain of potential correspondences of interest in advance, so as to allow application of statistical models to assess the likelihood that any set of observed correspondences would have occurred by chance.
- *Poor observation*. Sometimes our crucial observations are inaccurate (*e.g.* Wiseman and Morris, 1995a). Our senses were not working well or the conditions for observation were poor. Perhaps we did not know where or when to look. Or, our observations were only partial, as we did not realize the importance of additional observations. Thus we may miss a possible causal linkage between the two events.
- *Misunderstanding of observations*. Our observations may be accurate but we are not knowledgeable enough to interpret them. We do not know about certain electrostatic effects or what the heat of the hand can do to a stopped watch. We may not think to apply our knowledge to the present situation, not recognizing its potential applicability. Or, we may make simple reasoning errors in all innocence.
- *Poor storage and retrieval of memories*. Our memories may be poor when it comes to reconstruction of events, which at the time may not have seemed important to us. Problems can occur either during the consolidation and storage of information or during recall and recognition.
- *Self-deception*. If motivated, we may have a desire to observe evidence for or against the existence of certain effects. This may lead us to deceive ourselves by biasing our observational strategies, where and when we pay attention, and so on. We may have a selective memory favoring confirmatory events. We may engage in intentional misprocessing of information, self-deceptive analyses, and so on.
- *Deception by others*. Sometimes we can be victimized by deliberate tricksters motivated to produce false evidence. There is a sizeable literature on the fraudulent production of MMI effects. Frauds have devised many strategies for interfering with observations, for renegotiating the interpretations of events more favorably after the fact, even for deliberately producing patterns in the data of whatever sort might be of interest to the observers. Deception can occur in the stage of reporting observations as well. A recent guide to the literature on such strategies and how to deal with them can be found in Wiseman and Morris (1995b).

- *Biological distortion of processing*. The processing of information by our brains can be biased by internally and externally generated chemicals including diet, sleep conditions, disease, or traumatic brain damage. All can affect our sensing, perceiving, higher processing, storage and retrieval, as well as formation and implementation of intentions. In more extreme cases we may experience hallucinations, thought disorders, disorders of volition, and so on (see Neppe, 1993, for an overview of brain dysfunction, psychopathology, and anomalous experience). Such distortions may affect the experiences of an individual such as to make that person appear to themselves to be showing MMI effects; or they may lead an external observer to attribute MMI effects to others.
- *Functional distortions of processing*. Information processing can also be strongly distorted in individuals who have acquired a powerful belief system and have carried it to the point of a delusional system (*e.g.* Oltmans and Maher, 1988). Sometimes such distortion is enhanced in part by neurophysiological problems as in those mentioned above, which contribute to the formation of the beliefs, but then the beliefs and consequent distortion of processing can take on a life of their own.
- *Hidden natural causes*. Sometimes there can be causes in the system that are hidden from the observers such that these causes would not be evident through the usual strategies of selecting what observations to make. The observer may not know of the existence of certain system components that are in fact exerting an influence.
- *New application of existing principles*. As science progresses, we learn new facts and new principles, without necessarily learning all the ways that these principles can be applied. Some forms of animal communication were mysterious until they came to be understood as involving sonar, itself only recently understood at the time. Certain substances such as medicinal plants or metals may have properties that are presently not yet known but when understood will be seen to involve no truly new principles.
- *Presently unknown natural causes*. Under conditions apparently sufficient to rule out the above ten categories and all presently understood natural processes, do any effects remain? If so, can they be studied such as to develop an understanding of the new natural processes involved? This area is generally regarded as the province of experimental parapsychology and is where MMI research would also be located.
- *Causation beyond nature*. Is there anything left after the above, *e.g.* anything truly beyond the laws of nature as we will come to understand them? This is the realm of the so-called "supernatural." In principle there may well be, depending on how we define the laws of nature. In practice it may be impossible ever to do anything more than discover effects for which there are not yet known natural laws. And of course science continues to provide non-supernatural interpretations of effects

formerly regarded as potentially supernatural by some observers. Certain issues such as the "Hard Problem" of the mind-body debate (*e.g.* Chalmers, 1996) do remind us that we cannot yet demonstrate that consciousness itself (and associated concepts such as qualia and will) is determined solely by natural laws. At any given level of apparent lawfulness there may be hidden lawlessness lurking below; at any given level of apparent lawlessness there may be rich lawfulness hiding just out of sight. In general, the very amenability of an effect to experimentation serves as an indicator that it is at least in part lawful. The processes of systematic scientific research, by which effects are demonstrated, models developed and tested, aims to address natural causation and can have little to say about any truly supernatural causation, should any such exist.

A Communication Model for Designing MMI Experiments

Components of the Model

One of the most convenient ways to look at the diversity of events and experiences that have been interpreted as MMI effects is to insert them into a simple communication model (Morris, 1980). ESP becomes the flow of information or influence from a target source in the environment to the receiver organism, despite the presence of barriers between them that should prevent all currently understood channels of communication. PK becomes the flow of information or influence from a source organism to a target receiver in the environment. Constructing experimental designs to elicit and explore evidence

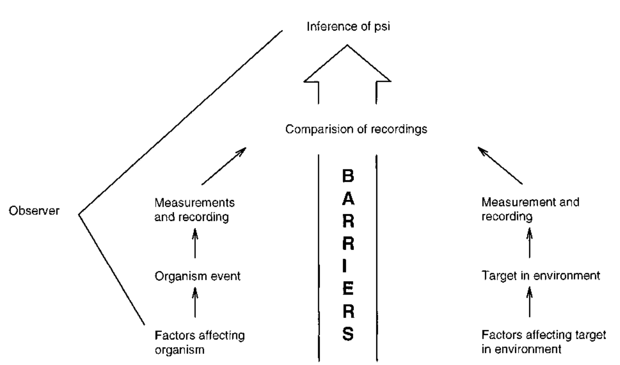

Fig. 1. A model of factors affecting an observer's assessment of the likelihood that MMI has occurred

Fig. 1. A model of factors affecting an observer's assessment of the likelihood that MMI has occurred.

for MMI effects then becomes a matter of understanding the various factors that can influence such communication systems and our observations of them. To organize our understanding of the many factors involved, we need a model of what goes on when we interpret a set of events as indicating MMI. Figure 1 illustrates such a model, modified from Morris (1986). It can be regarded as a system having various component subsystems that may contribute to an observer's assessment of whether or not MMI has taken place. It can be applied to either ESP or PK.

The *observer* is the person (not just a physical apparatus) who forms a judgment about whether or not MMI has taken place, on the basis of information derived from the other components of the system.

If the *organism* is being evaluated for ESP, then we attend to its experience as reflected in its behavior, or just to the behavior itself, to see if it appears to have received input from an external event. This may take the form of a behavioral choice among alternatives, a verbal report of one's experiences, a physiological response, and so on. Comparably if we are looking for evidence of PK, then we ask whether the organism's needs and intentions appear to have had influence upon some aspect of the environment. A gambler needs the dice to produce winning numbers; a healer wants clients to get well. The observer and the organism may be the same, as when an individual is attempting to evaluate whether they have been producing MMI effects.

The *target* is any aspect of the environment with which the organism appears to be interacting without access to presently understood means. For instance in ESP it can be a remote physical location or a hidden playing card; for PK it can be the fall of dice or a physiological response. In an experiment it is the aspect with which the organism has been requested to interact, or which is being monitored to see if spontaneous interactions appear to be taking place.

Barriers are the situational factors which would appear to prevent any and all presently understood means of interaction from taking place. In experiments they are generally put in place deliberately by the research team. They can include distance, time, or shielding of some sort.

If we are to understand the relationship between the organism event (the behavior or the needs and their associated intentions) and the target, we must understand as much as we can the *set of factors that might affect the characteristics of that organism event,* so as to assess whether or not any of those factors may bias the event to resemble the target.

Comparably, we must understand as much as we can about the *set of factors that might affect the characteristics of the target,* to assess whether they might bias the target to resemble the organism event.

There must be *measurement and recording of the organism event* in some objective way, such as not to be influenced by knowledge of the target or its associated systems.

Comparably, there must be *measurement and recording of the target event* objectively such as not to be influenced by knowledge of the organism event or its associated systems.

There must be a *comparison of recordings* at some stage to assess their degree of similarity using some specified criteria. Such comparison must either be on a pre-specified objective basis, or else evaluated subjectively by evaluators who are prevented from being influenced by their biases through some formal procedure such as blind judging, as described below.

Generally there is a *comparison with chance* using a mathematical model to estimate how often the obtained degree of similarity is likely to occur just by chance. Various statistical procedures can be used, as described below. The exceptions are cases in which the effect is so anomalous that in the opinion of the observer it is unlikely ever to occur by chance, such as certain large-scale PK phenomena.

Once the above steps have taken place, the outcome can be evaluated in terms of an *inference of MMI* by those sufficiently aware of the various components of the system and how they have interacted in the past that they can realistically assess the likelihood that the organism and target were linked through presently understood means.

Application of the Model

By bearing in mind these component subsystems and their interactions, we can now consider some of the ways that we can be misled in our attempts to interpret any specific evidence for MMI.

First, observers may have inadequate access to the other subsystems which they need to observe. They may be denied this access, deliberately or accidentally, because they do not know what to observe or when and where to observe. They may be given inaccurate information, deliberately or accidentally. They may misperceive the information. They may have their attention diverted. They may misinterpret or misremember the information. Secondary observers may also degrade the information, deliberately or accidentally.

Second, there may be too much interaction among subsystems *via* presently understood means. The two sets of antecedent factors, those affecting the organism and those affecting the target, may be interactive. One may influence the other or they may share too many common influences. The barriers may be less effective than anticipated. They may be weaker, they may be penetrated *via* trickery, or the target and organism subsystems may emit and receive information in more ways than realized. There may also be too much interaction among the two descriptive subsystems. Either may influence the other, or they may themselves share factors not prevented by any barriers. Assessment of the degree of correspondence may be biased unless a blind or objective comparison procedure is in place. Essentially, any interaction through conventional means between any pair of subsystems on opposite sides of the barrier can influence the assessed degree of correspondence. Finally, the evaluation of the role of chance must be done carefully, with statistical models appropriate to the design of the study and the measurements taken.

Third, there may be external systems which are capable of interacting with some or all of the subsystems described above. Essentially, any external system that interacts through conventional means with one or more subsystems on the organism side of the barrier and one or more subsystems on the target side of the barrier can influence the assessed degree of correspondence. Such an external system could be a researcher or confederate, or a common aspect of the environment such as themes in the news or even the weather.

Adequate experimental design must take all of the above into account. Many of the points made above are simply good normal scientific practice, but such good practice can be difficult to apply in investigating anomalous effects within complicated systems, a characteristic of most if not all MMI research. There is no such thing as the perfect experiment although we can approximate it. There are certain criticisms that can always be put forth and are essentially unfalsifiable: undetected participant fraud; undetected artifact; improper implementation of procedure; inadequate description of the procedures used; experimenter or investigator fraud; and chance (unless the evidence in question is so anomalous that it would not have occurred at all given our present scientific understanding). Such issues become of less concern once one is dealing with effects that are readily obtainable under a variety of circumstances by many researchers, and once one is involved in doing systematic model testing.

Translation into Experimental Designs

ESP Designs

An overview of ESP research designs can be found in Morris (1978, 1982).

Restricted choice designs evaluate the ability to be a good chooser from amongst known alternatives. The targets are a series of objects or events drawn from a set known to the receiver, such as the suits in a deck of cards or a set of marked positions corresponding to the possible positions of hidden objects. The order of the targets must be sufficiently randomized to eliminate its predictability or its similarity to any pattern of response by the receiver. The receiver knows the possibilities and makes a series of responses (like guesses) which can then be compared objectively with the order of the actual targets to assess whether the number of actual correspondences exceeds that expected by chance to a meaningful extent. For an extended discussion of methodological issues regarding such procedures, see Rhine *et al*. (1966).

Free response designs evaluate the ability to respond in some manner to a qualitatively rich target such as a picture, film, or complex object. These targets are less restricted in content and form, such that the receiver produces impressions rather than choices, at least initially. The target must still be selected randomly from among possible options so that the receiver cannot infer the target characteristics. Evaluation of correspondence is more complicated and must be done blind, linked with whatever model is used for analysis. Three examples will illustrate. First, one can assign categories to the target material, as in the Maimonides target slide pool (Honorton, 1975), which was a set of slides representing all possible combinations of the presence and absence of ten content categories; or the coded set of target location pictures developed by Jahn, Dunne and Jahn (1980). In both cases the impressions of the receiver could then be coded along similar dimensions and the number of correspondences compared against chance, making sure to take receiver biases toward certain categories into account. Second, one could start with a pool of targets preselected to be unlike each other such that one such pool is randomly selected from among available pools, then within that pool a target is randomly selected. The receiver produces a set of impressions: oral, written, or drawn, which are then shown to a blind judge (in some designs the receiver serves also as the judge) who must compare the description with each of the target possibilities in the pool. The judge may be asked to rate all on a scale of resemblance, or to rank each of the potential targets from most to least resemblance. The ratings or rankings assigned to the correct targets are then compared to those assigned to the non-targets to see if the difference in ratings exceeds that which would be expected by chance (*e.g.* Honorton *et al.*, 1990). A third strategy is to have raters assign degree of correspondence ratings to the targets and corresponding impressions themselves, without controls, while remaining blind as to which targets were from sessions of one kind as opposed to another kind. This procedure can be used when comparing success rate between conditions, not as an absolute measure of success, given that raters are always aware that the target for comparison is in fact the correct one. Examples of free response designs include the ganzfeld studies in which concealed pictures or film clips serve as targets (*e.g.* Bem and Honorton, 1994) or the remote viewing studies in which targets were geographical locations (*e.g.* Puthoff and Targ, 1976).

In *somatic designs*, the target has some property expected to provoke a shift in somatic activity in the receiver. The relevant somatic system is monitored and changes in its activity over time provide the measurement to be correlated with the timing of the remote target event. In one kind of design, shifts in the arousal of an agent or sender are manipulated, to see if they correlate with shifts in the receiver's arousal. In another design the receiver is suddenly presented with high or low arousal material, or startled in some way and receiver arousal monitored throughout to see if there is any anticipatory shift in arousal just prior to the stimulus onset, but before the receiver would have had conventional access to indicators that the stimulation was about to occur (*e.g.* Bierman and Radin, 1997). In either design, somatic measures are taken under conditions in which they are expected to be differentially active, such as high and low conditions, stimulus *vs.* control period, and so on. Standard precautions against artifacts such as drift in arousal during a session must be taken into account through counterbalancing of conditions or through the processing of the data itself. Appropriate statistical processing of the data can be done in many ways. Successive measures taken from somatic systems can be heavily intercorrelated and this must be taken into account when defining units of analysis.

PK Designs

An overview of PK research designs can be found in Rush (1977, 1982).

In *discrete outcome designs*, the organism is designated as an agent, a source of influence or information, and asked to affect mentally an event which has a discrete set of possible outcomes, such as the fall of a die or the output of a source of randomness linked to a display. For instance, the agent is given a set of assigned outcomes, such as a counterbalanced or randomized order of the digits one through six, and then asked to influence the consecutive mechanical throws of dice to conform to the assigned instructions. At the end of a session the number of occasions in which there was a correspondence can be compared to that expected by chance, or to a control condition (Radin and Ferrari, 1991). A similar design involves the use of sources of noise which produce an output which can then be processed such as to produce a random bit stream, which in turn can be used to drive a numeric or graphic display of some sort (Radin and Nelson, 1989; Schmidt, 1987). The agent (sometimes called an operator in MMI research) is given instructions to bias the observed display in different directions, such as the high aim *vs.* low aim *vs.* baseline triad of conditions used in the PEAR work (*e.g.* Jahn *et al.*, 1997). Or they can attempt to influence a property of a non-visual display, such as pitch of a tone or click rate (*e.g.* Schmidt, 1973). In each case the instructions are compared with the observed outcomes and assessed with respect to the degree of resemblance expected by chance. Some designs use large numbers of bits per trial, summed in some way to produce a single shift in the display, thus allowing the evaluation of more complex patterns in the data and the testing of models about the nature of the effect. These procedures are among the most theoretically useful in MMI. Such designs illustrate that often more than one supposed psi or MMI ability may be working in the same experiment. For instance, the agent may be assigned a fixed instruction for each trial; or they may be given some freedom to choose the target condition from trial to trial, so long as proper counterbalancing is maintained. Under the latter circumstances one might argue that the results are interpretable as precognition just as easily as PK and some theorists (*e.g.* May, Utts and Spottiswoode, 1995) have argued that all discrete outcome studies are in principle interpretable within an ESP or anomalous cognition (their preferred term) framework. Additionally, it could be argued that such PK influences could occur in any process of target selection which incorporates a source of randomness, leading for instance to the selection of favorable targets in a free response procedure. Sometimes such procedures are said to be tests for “micro-PK” because in physical terms the effects appear very small.

Stable system designs involve attempts by an agent to influence the output of some sort of stabilized system, which has measurable fluctuations. One example would be the use of a partially closed physical system such as a thermistor in a sealed container, known to have low level fluctuations with a degree of randomness to them. An agent would then be given a set of instructions to increase or decrease the recorded temperature in either a randomized or counterbalanced order, during a set of discrete epochs in the course of a session (*e.g.* Schmeidler, 1973). A second example procedurally resembles the somatic designs described above. An agent is asked to influence the level of activity of a biological system such as a rodent in a running wheel, the hemolysis rate of a blood sample, or the physiologically monitored level of arousal in a human receiver. For the last, the design can be identical to the somatic design, but with the emphasis placed on the active role of the agent rather than the receiver. As above, there are assigned instructions to increase and decrease activity, plus no influence control periods. These designs are often now referred to as DMILS (Direct Mental Interaction with Living Systems) procedures (see Schlitz and Braud, 1997, for an overview). In some DMILS procedures involving two people, when each is aware of and interested in the intentions of the study, either or both could be actively involved; a successful session might be regarded as ESP or PK, anomalous cognition or anomalous perturbation, and so on.

Static system designs involve attempts by the agent to induce movement or changes in movement of a static system. These designs vary according to the nature of the target system and basically involve isolating the system from known means of influence, then monitoring it to see if any movements occur when an agent is attempting to induce such movements (*e.g.* Rush, 1977, 1982). Statistical inference is less likely to be involved in these cases, as the test system is set up such that any measured movement would itself be anomalous and, if correlated with an agent's apparent intentions, indicative of an MMI effect. Sometimes these procedures are regarded as involving "macro-PK" because the effects are physically relatively large, obvious to the naked eye in a given instance. Much less formal work has been done with such designs and the database is small. One problem is that such effects are much more the kinds of effects produced by magicians' strategies, which must be thoroughly understood and guarded against when setting up the target system and the measurement system as well (Hansen, 1990; Wiseman and Morris, 1994). It is thus especially important that static system designs have input from magicians with relevant expertise, although magical expertise is very useful for the other designs also, especially when working with individual claimants. Truzzi (1997) provides an excellent overview of the many issues involved in the use of magicians in parapsychological/MMI research designs.

Interpretation of Results

Most researchers set up their designs in such a way that they are looking for evidence of a causal linkage between organism mentation (mind) and target in the environment (matter). The communication model presented above illustrates this. However, the existence of strong correspondence still leaves open a host of other interpretations as well.

Recall this guideline: *Essentially, any interaction through conventional means between any pair of subsystems on opposite sides of the barrier can influence the assessed degree of correspondence*. By similar reasoning, any interaction through non-conventional means between any pair of subsystems on opposite sides of the barrier could also influence the assessed degree of correspondence. This opens the door to a variety of additional subcomponent interactions beyond those which may have been intended by the research team. In any experimental system there may have been many different MMI effects taking place to produce a given degree of correspondence, as has been noted in detail elsewhere (*e.g.* Morris, 1975, 1980).

Comparable complications arise when considering the second guideline: *Essentially, any external system that interacts through conventional means with one or more subsystems on the organism side of the barrier and one or more subsystems on the target side of the barrier can influence the assessed degree of correspondence*. Any external system interacting through non-conventional means with subsystems on both the organism and target sides of the barrier could influence the assessed degree of correspondence as well. This external system could be a third party producing an ESP or PK interaction with the subsystems (*e.g.* Honorton, 1976); or, it could be some situational or environmental factor such as those construed by Jung as responsible for synchronistic events (*e.g.* Bender, 1977). A related possibility is an external system interacting conventionally with one side and non-conventionally with the other. For these reasons, it could be argued that the original communication model offered above is too simplistic and should be expanded to acknowledge the diversity of interactions possible in any given experimental system using any of the above kinds of designs. Some, such as Stanford (1978, 1981), have argued that we are better off avoiding even the implication of communication and instead adopting a Conformance Behavior Model, which simply states that observed correspondences in parapsychology studies can be taken to represent a conformance between target and organism and no more at this stage. Such an approach does not forbid a process oriented research program in which specific models can be developed, refined, and tested empirically. For an overview of these and other theoretical approaches within parapsychology, see Stokes (1987).

Some Measurement Issues in Mind-Matter Research

Evaluation of the Null Hypothesis

Most of the conceptual issues involved in MMI measurement are the same as those in any other area of research, about which much has been written. In the remainder of this paper I would like to focus on some related issues which have emerged within MMI research due to some of its unusual properties. A basic strategy mentioned throughout the design descriptions covered above is the use of statistical tests to compare the degree of similarity between the target and organism descriptive systems with that which would be expected by chance. This has essentially been regarded as a test of what has been called the "null hypothesis," the hypothesis that any observed differences are not real effects, but are within the range of those expected by chance. We take a model of what would happen just by chance, generally a known distribution; we take the observed outcome; and we calculate how often such an observed outcome would occur just by chance alone.

This may produce a statement of the sort: The likelihood of *X* outcome or more extreme occurring by chance alone is $p < 0.01$, therefore we reject the null hypothesis. We must be careful in doing so, because the likelihood of *X* occurring just by chance is not necessarily the likelihood that *X* is not due to chance. In practice we tend to say, such outcomes rarely occur by chance, and the less likely they are to occur by chance the more willing we are to assume chance played less of a role. This is not unreasonable but we should remember that the precision of the *p*-value we obtain is not to be transferred to the question of real interest, namely how likely was the observed result to be due to a real effect.

An excellent coverage of these and related issues can be found in Gigerenzer *et al.* (1989). They contrast the Fisherian approach described above, based on comparisons of an observed outcome with that predicted under the null hypothesis, with the Neyman-Pearson approach of comparing two rival hypotheses, of testing one against the other, specifying the statistical outcome conditions under which one would be preferred over the other.

Comparison of observed departures from chance with models of chance distributions is used in evaluating the overall strength of an effect, for comparison of outcomes among different conditions, for evaluating the strength of correlations with other variables, and for the assessment of patterns within the data structure. For the reasons mentioned above and more, the assessment of the strength of an effect is most suitably calculated (or estimated) as an effect size, using any of several standard equations for doing so (*e.g.* Rosenthal and Rosnow, 1991). The effect size can then be taken into account in estimating the power of any given study to detect such an effect to a significant extent (Cohen, 1977).

The Problem of Psi Missing

In MMI research, we have the concept of psi missing, apparently meaningful negative outcomes of the sort that would occur by chance significantly rarely. These effects have seemed meaningful because they have often been consistently correlated with other variables that make sense, such as the "sheep-goat effect," the finding that those with negative attitudes toward MMI will score below what is expected by chance (Palmer, 1977), or the various studies of negative results during induced negative experimenter effects (*e.g.* Honorton, Ramsey and Cabbibo, 1975). Such negative departures have thus come to be seen as evidence themselves for MMI effects. Conceptually psi missing has been construed either as motivated miscommunication as in the sheep-goat effect, where participants produce the results they regard as confirming their views (Lovitts, 1981), or as a kind of "wires crossed" misprocessing, *e.g.* by consistent symbol confusions (Kelly *et al.*, 1975). But taking psi missing seriously opens up some additional lines of reasoning.

Ordinarily when comparing outcomes with a chance distribution we specify what regions of that distribution are of interest to us. If I have no interest in a negative outcome, if I am prepared to regard any outcome in that direction as simply a chance fluctuation, then I will use a one tail test, meaning that I have increased power to detect my effect if it is there, but in exchange I must ignore any outcome in the other direction, no matter how strong. If we are not careful to specify the number of tails in advance, we can easily make an after the fact decision, effectively allowing ourselves an extra opportunity to declare a significant outcome. This problem has been addressed by some by just reporting effect sizes and related two tailed p-values, then allowing readers to draw their own inferences.

However another problem is that we now have two kinds of measures of MMI, positive deviations and negative deviations. Do they both mean the same thing? They do not behave similarly in their relationship to other variables, and they would seem to involve somewhat different processing. One can easily erect problematic examples in which, for instance, there is a strong negative effect size under condition A and a weak positive effect under condition B. Assuming this differential is replicable, which condition is more conducive to MMI effects?

If departures from chance can occur meaningfully in both directions, then an additional strategy is to adopt variance measures in addition to or instead of the success rates themselves. Once again, we can compare variance measures with chance distributions and obtain variances that would occur by chance very rarely, being either significantly high or significantly low. The latter would mean that the obtained set of scores was consistently closer to chance than would be expected by chance. Both kinds of variance results have already been found in the research literature (*e.g.* Stanford, 1966; Whittlesey, 1960). By adding in this measure of MMI effects we further complicate the picture, however, and can continue to construct troublesome examples in which condition A has good results by one measure and poor results by another, with the reverse true for condition B, and so on.

Do we conclude that both are producing MMI effects, that each produces a specific kind of MMI effect with different processing involved; or do we regard one kind of departure from chance as more meaningful than others? For practical applications it would seem best to focus on psi-hitting, but in general we would want to understand the full range of MMI manifestations. And even in practical terms the presence of highly complex patterns in databases, if they can be identified as specific to a particular participant or experimenter, may provide us with a valuable tool to identify which organism(s) are involved in the data production, thus addressing the experimenter psi aspect of the "Experimenter Effect" (Palmer, 1997; Schmeidler, 1997) very empirically.

Although psi hitting is how we have traditionally and pragmatically defined success in communication, there have been attempts to combine several indicators through the use of samples predicting to the main data. Carpenter (1983) used samples of data generated within a session to determine empirically whether the sample showed psi hitting or missing, combined with mood indicators predictive of large or small variance, and then applied the appropriate test to the remaining data for the session, producing strong positive results but with a large investment of time in generating the data. Too little work has been done in this area. The techniques are easiest to apply in restricted choice and discrete outcome studies and harder with analogue measures, but could still be done. In addition to basic hitting rate, sample based predictions could also be applied to variances in the values of correlations, differences among conditions, complex internal patterns in the data, and so on. Burdick and Kelly (1977) provide an extensive review of the main statistical methods used in parapsychology and the many problems associated with their application to the experimental database, along with recommendations for improvement.

Conclusion

The bottom line is that in MMI research, we use evaluation of overall resemblance between organism measure and target measure to estimate size of effect; and evaluation of differential resemblance to understand the underlying processes and build models, just as we would with source and receiver measurements in any other study of communication. The creative energy of researchers has largely gone into identifying the basic research strategies needed to rule out various alternative interpretations that have been offered by critics for the MMI evidence; or, for many, attempting to demonstrate the reality of MMI by statistical analyses which can be seriously questioned when used in that context. Statistical evidence in itself is never sufficient as a contribution to the corpus of scientific knowledge; it must be linked to a testable model or even theory of some sort. Most research to date has involved relatively gross or *ad hoc* measurement and unsystematic analysis techniques, not designed to extract the maximum information available to characterize degree of absolute or differential resemblance. This situation is starting to change, but only to the extent that we move on to the stage of understanding and not just demonstrating. Then we will be clearly beyond the stage of appearing to be studying a set of effects having only negative definitions.

References

Bem, D. and Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, 115, 4–18.

Bender, H. (1977). Meaningful coincidence in the light of Jung-Pauli's theory of synchronicity. In B. Shapin and L. Coly (Eds.), *The Philosophy of Parapsychology*. New York: Parapsychology Foundation, 66–85.

Bierman, D. and Radin, D. I. (1997). Anomalous anticipatory response on randomized future conditions. *Perceptual and Motor Skills*, 84, 689–690.

Burdick, D. and Kelly, E. (1977). Statistical methods in parapsychological research. In B. Wolman (Ed.), *Handbook of Parapsychology*. New York: Van Nostrand Reinhold Company.

Carpenter, J. (1983). Prediction of forced-choice ESP performance. Part II. Application of a mood scale to a repeated guessing technique. *Journal of Parapsychology*, 47, 217–236.

Chalmers, D. (1996). *The Conscious Mind*. New York: Oxford University Press.

Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences*. 2nd edition. New York: Academic Press.

Diaconis, P. and Mosteller, F. (1987). Methods for studying coincidences. *Journal of the American Statistical Association*, 884, 853–861.

Gigerenzer, G., Swistink, Z., Porter, T., Daston, L., Beatty, J. and Krüger, L. (1989). *The Empire of Chance*. New York: Cambridge University Press.

Hansen, G. P. (1990). Deception by subjects in psi research. *Journal of the American Society for Psychical Research*, 84, 25–80.

Honorton, C. (1975). Objective determination of information rate in psi tasks with pictorial stimuli. *Journal of the American Society for Psychical Research*, 69, 353–359.

Honorton, C. (1976). Has science developed the competence to confront claims of the paranormal? *Research in Parapsychology,* 1975, 199–224.

Honorton, C., Ramsay, M. and Cabibbo, C. (1975). Experimenter effects in extrasensory perception. *Journal of the American Society for Psychical Research*, 69, 135–149.

Honorton, C., Berger, R. E., Varvoglis, M. P., Quant, M., Derr, P., Schechter, E. I., and Ferrari, D. C. (1990). Psi communication in the ganzfeld: Experiments with an automated testing system and a comparison with a meta-analysis of earlier studieFs. *Journal of Parapsychology*, 54, 99–139.

Jahn, R. G., Dunne, B. J. and Jahn, E. G. (1980). Analytical judging procedure for remote perception experiments. *Journal of Parapsychology*, 44, 207–231.

Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H., and Bradish, G. J. (1997). Correlations of random binary sequences with pre-stated operator intentions: A review of a twelve-year program. *Journal of Scientific Exploration*, 11, 345–367.

Kelly, E., Kanthamani, H., Child, I. L., and Young, F. W. (1975). On the relation between visual and ESP confusion structures in an exceptional ESP subject. *Journal of the American Society for Psychical Research*, 69, 1–32.

Lovitts, B. E. (1981). The sheep-goat effect turned upside down. *Journal of Parapsychology*, 45, 293–309.

May, E., Utts, J. and Spottiswoode, S. J. (1995). Decision augmentation theory: Toward a model of anomalous mental phenomena. *Journal of Parapsychology*, 59, 196–220.

Morris, R. L. (1975). Building experimental models. *Journal of Communication*, 25, 117–125.

Morris, R. L. (1978). A survey of methods and issues in ESP research. In S. Krippner (Ed.), *Advances in Parapsychological Research 2*. New York: Plenum Press, 7–58.

Morris, R. L. (1980). Psi functioning within a simple communication model. In B. Shapin and L. Coly (Eds.), *Communication and Parapsychology*. New York: Parapsychology Foundation, 1–24.

Morris, R. L. (1982). An updated survey of methods and issues in ESP research. In S. Krippner (Ed.), *Advances in Parapsychological Research 3*. New York: Plenum Press, 5–40.

Morris, R. L. (1986). What psi is not: The necessity for experiments. In H. L. Edge *et.al.*, *Foundations of Parapsychology*. Boston: Routledge and Kegan Paul, 70–110.

Morris, R. L. (1996). Koestler's legacy: Parapsychology at the University of Edinburgh. *The Skeptic*, 10 (5, 6), 6–8.

Neppe, V. (1993). Clinical psychiatry, psychopharmacology and anomalous experience. In L. Coly and J. McMahon (Eds.), *Psi and Clinical Practice*. New York: Parapsychology Foundation, 145–162.

Oltmans, T. and Maher, B. (Eds.) (1988). *Delusional Beliefs*. Chichester: John Wiley and Sons.

Palmer, J. (1977). Attitude and personality traits in experimental ESP research. In B. B. Wolman (Ed.), *Handbook of Parapsychology*. New York: Van Nostrand Reinhold, 175–201.

Palmer, J. (1997). The challenge of experimenter psi. *European Journal of Parapsychology*, 13, 110–125.
Puthoff, H. E. and Targ, R. (1976). A perceptual channel for information transfer over kilometer distances: Historical perspectives and recent research. *Journal of the Institute of Electrical and Electronic Engineers*, 64, 329–354.
Radin, D. I. and Ferrari, D. C. (1991). Effects of consciousness on the fall of dice: A meta-analysis. *Journal of Scientific Exploration*, 5, 61–84.
Radin, D. I. and Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics*, 19, 1499–1514.
Rhine, J. B., Pratt, J. G., Smith, B. M., Stuart, C. E., and Greenwood, J. A. (1966). *Extrasensory Perception after Sixty Years*. Boston: Bruce Humphreys (originally published 1940).
Rosenthal, R. and Rosnow, R. (1991). *Essentials of Behavioral Research: Methods and Data Analysis*. New York: McGraw Hill, 440–451.
Rush, J. (1977). Problems and methods in psychokinesis research. In S. Krippner (Ed.), *Advances in Parapsychological Research 1*. New York: Plenum Press, 15–79.
Rush, J. (1982). Problems and methods in psychokinesis research. In S. Krippner (Ed.), *Advances in Parapsychological Research 3*. New York: Plenum Press, 83–114.
Schlitz, M. and Braud, W. (1997). Distant intentionality and healing: Assessing the evidence. *Alternative Therapies*, 3, 62–73.
Schmeidler, G. R. (1973). PK effects upon continuously recorded temperature. *Journal of the American Society for Psychical Research*, 67, 325–340.
Schmeidler, G. R. (1997). Psi-conducive experimenters and psi-permissive ones. *European Journal of Parapsychology*, 13, 83–94.
Schmidt, H. (1973). PK tests with a high speed random number generator. *Journal of Parapsychology*, 37, 105–118.
Schmidt, H. (1987). The strange properties of psychokinesis. *Journal of Scientific Exploration*, 1, 103–118.
Stanford, R. G.(1966). A study of the cause of low run-score variance. *Journal of Parapsychology*, 30, 236–242.
Stanford, R. G. (1978). Toward reinterpreting psi events. *Journal of the American Society for Psychical research*, 72, 197–214.
Stanford, R. G. (1981). Are we shamans or scientists? *Journal of the American Society for Psychical research*, 75, 61–70.
Stokes, D. (1987). Theoretical Parapsychology. In S. Krippner (Ed.), *Advances in Parapasychology 5*. Jefferson, NC: McFarland Press, 77–189.
Truzzi, M. (1997). Reflections on the sociology and social psychology of conjurors and their relations with psychical research. In S. Krippner (Ed.), *Advances in Parapsychological Research 8*. Jefferson, NC: McFarland Press, 221–271.
Whittlesey, J. R. B. (1960). Some curious ESP effects in terms of variance. *Journal of Parapsychology*, 24, 220–222.
Wiseman, R. and Morris, R. L. (1994). Modelling the stratagems of psychic fraud. *European Journal of Parapsychology*, 10, 31–44.
Wiseman, R. and Morris, R. L. (1995a). Recalling pseudo-psychic demonstrations. *British Journal of Psychology*, 86, 113–125.
Wiseman, R. and Morris, R. L. (1995b). *Guidelines for Testing Psychic Claimants*. Amherst, N. Y.: Prometheus Press.