8305349

Kern, Leslie Helen

THE EFFECT OF DATA ERROR IN INDUCING CONFIRMATORY INFERENCE STRATEGIES IN SCIENTIFIC HYPOTHESIS TESTING

The Ohio State University PH.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark _✓_.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages _✓_
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _✓_
11. Page(s) __________ lacking when material received, and not available from school or author.
12. Page(s) __________ seem to be missing in numbering only as text follows.
13. Two pages numbered __________. Text follows.
14. Curling and wrinkled pages _____
15. Other ____

THE EFFECT OF DATA ERROR IN INDUCING CONFIRMATORY INFERENCE STRATEGIES IN SCIENTIFIC HYPOTHESIS TESTING

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Leslie Helen Kern, B.A., M.A.

* * * * *

The Ohio State University

1982

Reading Committee:

Michael E. Doherty
Herbert L. Mirels
Thomas E. Nygren
Richard K. Russell

Approved By

Adviser
Department of Psychology

ACKNOWLEDGEMENTS

More than anyone else who helped me throughout the writing of this dissertation, I want to thank my brother, Hadley, who spent hundreds of hours turning my research ideas into the polished computer programs that made this study possible. He never just did what I asked; he always refined and improved my initial proposals, so that the ultimate design was simpler and more efficient than it ever could have been without him. The debt I owe to him is beyond repayment.

A number of others made significant contributions to this project. Had I not been eavesdropping on a conversation between Jack Mynatt and Hillel Einhorn at an MPA meeting last spring, the significance of the questions addressed in this dissertation would never have occurred to me. Moreover, subsequent conversations with Jack led to the development of an experimental methodology which was, if I do say so myself, rather clever. Mike Doherty helped me to refine and simplify the initial design so that it was possible to complete the research in less than a decade. Moreover, he was enormously generous with his time, spending more of it editing and talking with me than I might have expected had I been a student at his own university. Herb Mirels, aside from being one of the warmest and most supportive persons I have encountered over the years in graduate school, consistently took a genuine interest in my research despite the fact that it had no relevance whatsoever to implicit personality theory or the Protestant ethic. My discussions with Rich Russell were invaluable in helping me to see the implications of my study for the practice of clinical psychology, which is, after all, a major interest of mine. And Tom Nygren always responded with grace to those phone calls which are the bane of a statistician's existence--those that invariably begin with, "I have just one quick question..."

My father and mother, Morris and Xarifa Bean, worked very hard for many years in order that their children could have as much education as they wanted. For that, I am more grateful than I can express in words. The burned dinners and faded laundry I've received from my husband, Richard, over the past few months made me know I was loved, despite the fact that writing a dissertation has made me, at times, rather unlovable. Finally, were Patti Watson not the most competent, unruffled, and rapid typist in the world, this dissertation would never have been received by any of my committee members on time.

Thank you all for your enormous support and encouragement!

VITA

January 18, 1955.	Born - Springfield, Ohio
1972.	High School Diploma, Yellow Spring High School, Yellow Springs, Ohio
1972-1973	Research Assistant, Growth and Genetics Department, Fels Research Institute, Yellow Springs, Ohio.
1977.	B.A., University of Cincinnati, Cincinnati, Ohio.
1977-1978	University Fellow, Graduate Study in Clinical Psychology, The Ohio State University, Columbus, Ohio.
1978-1979	Teaching Assistant, Department of Psychology, The Ohio State University, Columbus, Ohio.
1979.	M.A., Clinical Psychology, The Ohio State University, Columbus, Ohio.
1979-1980	Teaching Associate, Department of Psychology, The Ohio State University, Columbus, Ohio.
1980-1982	Doctoral candidate in Clinical Psychology, The Ohio State University, Columbus, Ohio.

PAPERS

Warm, J.S., Seeman, W., Bean, L.H., & Chin, N. Meditation and sustained attention. Paper presented at the Annual Meeting of the Psychonomic Society. Washington, D.C., November, 1977.

PUBLICATIONS

Bean, L.H. Effects of the social circumstances surrounding the rating situation on students' evaluation of faculty. Teaching of Psychology, 1978, 5 (4), 200-202.

Kern, L.H., & Doherty, M.E. 'Pseudodiagnosticity' in an Idealized Medical Problem-Solving Environment. Journal of Medical Education, 1982, 57, 100-104.

Kern, L.H., Mirels, H., & Hinshaw, V. Scientists' understanding of propositional logic: An experimental investigation. Social Studies of Science, 1982, in press.

TABLE OF CONTENTS

Page

Page

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Despite the fact that science has for several centuries been a major social and political force influencing the beliefs and mental operations of nearly all persons, it is only very recently that the methods of science have been turned toward the investigation of science itself. Systematic examination of the cognitive processes involved in scientific thinking represents a new domain of inquiry within psychology, a domain of inquiry which regards science as an objectifiable phenomenon as worthy of empirical investigation as any other phenomenon within the natural universe.

Prior to the emergence of psychological studies of the cognitive mechanisms underlying scientific thinking, the activities of scientists had been the subject of extensive analyses by philosophers of science. Such analyses were typically of a formal nature, directed toward establishment of the logical relation between a scientific theory and a particular body of empirical data. Frequently, such analyses have demonstrated certain inferences made by scientists to be logically invalid (Popper, 1959), but have had questionable impact on the activities of practicing researchers. In fact, the majority of scientists pay no particular attention to the exhortations of philosophers of science regarding normative prescriptions for the conduct of research (Mitroff, 1974).

The negligible impact of formally derived "rules" for the conduct of scientific inquiry may be largely attributable to the fact that such prescriptions are not formulated in light of knowledge about the nature of scientific research as it is actually practiced.

The present investigation examines the behavior of individuals engaged in scientific hypothesis testing in an artificial environment designed to replicate many of the conditions under which actual research is undertaken. By manipulating certain parameters of this environment it is hoped that insight can be gained into the cognitive biases and limitations which may guide and/or constrain the activities of practicing scientists confronted with similar task dimensions in real-world empirical research. The results of the present study, to the extent that they increase understanding of science <u>as it is practiced</u>, may serve to complement the prescriptive norms formulated by philosophers of science, particularly regarding (1) how evidence should be related to hypotheses and (2) under what conditions hypotheses should or should not be abandoned.

CHAPTER II

The Bias to Confirm: Misleading Heuristic or Rational Adaptation to Ambiguity?

The Fallibility of Intuitive Inference

Empirical investigations of human judgment and decision making reveal numerous situations in which individuals fail to follow normative prescriptions of formal or statistical models in attempting to understand and predict events in the world around them. A decade of research shows that these errors in inference are systematic, reproducible and consistent across a variety of judgment tasks. Without attempting a comprehensive review of this literature, a few examples of recent investigations provide striking illustration of the fallibility of intuitive inference.

One well-documented shortcoming of intuitive judgment is the tendency toward overconfidence in one's conclusions. People are frequently wrong when they are <u>certain</u> that they are correct. Whether subjects are required to estimate the likelihood that they have responded correctly to a general knowledge question (Fischoff, Slovic, & Lichtenstein, 1977) or to assess the probability that a person practices a certain profession based on a scanty, unreliable personality description (Tversky & Kahneman, 1974), they express unwarranted confidence in the correctness of their judgments. Experienced clinical psychologists are not exempt from the tendency toward extreme confidence. Oskamp (1965) demonstrated that as the amount of information about a published case study increased, clinicians' confidence in their judgment acumen increased correspondingly, although the actual accuracy of their predictions remained unchanged.

Another frequently demonstrated failing of intuitive inference involves individuals' failure to follow Bayes' theorem when making inference judgments in sequential decision tasks (Shanteau, 1970, 1972). Ss are presented hypotheses and asked to determine which is correct based on the serial presentation of information. Following the presentation of each item of information, S assesses the likelihood that a particular hypothesis is true. These estimates, when compared with those specified by Bayes' theorem, show a tendency toward conservatism (Phillips & Edwards, 1966; cited in Pitz, Downing, & Reinhold, 1967). Even more remarkable, when Ss in such tasks receive information which is nondiagnostic (which has no relevance to judgments) they tend to make even less extreme inferences than when the diagnostic data is presented alone (Shanteau, 1975; Troutman & Shanteau, 1977; Tversky & Kahneman, 1974). Subsequent research has demonstrated that diagnostically worthless data dramatically affect likelihood estimates (Nisbett & Zukier, 1979; Zukier, 1979; Nisbett & Lemley, 1979; cited in Nisbett & Ross, 1980; Doherty, Mynatt, Tweney, & Schiavo, 1979) and attitude judgments (Youngblood & Himmelfarb, 1972).

Tversky and Kahneman (1971, 1973, 1974) and Kahneman and Tversky (1972, 1973) have demonstrated that individuals assessing the likelihood of an event or predicting the value of a quantity rely on judgment heuristics which reduce complex tasks to primitive, intuitive operations. Reliance on such heuristics often leads to successful inference, but at other times it results in serious, systematic error.

Numerous comparisons of the efficacy of clinical versus statistical modes of data collection and combination in psychotherapeutic settings reveal that mechanical modes have consistently superior predictive capability (Meehl, 1954; Sawyer, 1966). Moreover, clinical judgment inaccuracy appears very difficult to eradicate. Providing clinicians with extensive feedback concerning the accuracy of their inferences does not improve their judgment (Goldberg & Rorer, 1965; cited in Goldberg, 1968).

In short, research on human information processing suggests that certain cognitive biases promote systematic errors in inference across a variety of judgment tasks. Despite individuals' great confidence in the accuracy of their reasoning, considerable empirical evidence indicates that intuitive judgment can be highly fallible. The present investigation examines cognitive biases in inference which occur when individuals arrive at judgments through use of the hypothetico-deductive (H-D) method. Both the practicing psychotherapist attempting to discover lawful patterns in the behavior of another, and the professional scientist searching for functional relations between quantities, are attempting to evaluate the truth status of a particular hypothesis in light of relevant empirical data. With respect to application of the H-D method, a particular tendency which has come to be called "confirmatory bias" has such a pervasive and robust impact on inferential reasoning that a number of empirical studies have recently emerged which deal exclusively with this phenomenon. Confirmatory bias may manifest itself in any one of the following four ways (adopted from Mynatt, Doherty & Tweney [1978b], p. 4):

1. As a failure to seek evidence which might falsify a favored hypothesis.
2. As a failure to abandon a favored hypothesis once it is falsified.
3. As a failure to generate and test alternatives to a favored hypothesis.
4. As a failure to consider whether evidence supporting a favored hypothesis supports alternative hypotheses as well.

Research on Confirmation Bias

Wason (1960) demonstrated individuals' preferences for confirmatory over disconfirmatory reasoning by means of a task in which a number of incorrect but plausible hypotheses would readily suggest themselves to subjects. Because the possible instances of such incorrect hypotheses were, in principle, infinite, the task was such that the use of confirmatory reasoning alone would almost certainly lead subjects to erroneous conclusions. Participants were told that a triad of three numbers (2, 4, 6) conformed to a simple relational rule. They were asked to discover this rule by generating successive sets of three numbers and then inferring the rule from information given after each set (following the generation of a triad subjects were told whether or not it conformed to the rule). If an announced hypothesis was incorrect, participants were informed of this and then instructed to continue their search for the correct rule.

The rule to be attained was simply "three numbers in increasing order of magnitude." An erroneous hypothesis could be definitively disproved through the use of disconfirmatory reasoning (the generation of a triad that would yield "negative results" or information that the triad did not conform to the rule). However, participants could generate an infinite number of triads which would "confirm" such erroneous hypotheses as "even numbers increasing by twos."

Wason found that the typical subject attempted to solve the problem by generating a large amount of confirmatory "data" (triads which were implied by an [usually] erroneous hypothesis, i.e., 8, 10, 12; 20, 22, 24; etc.). Following the generation of a number of such triads, the subject would announce his hypothesis with certainty, and would be surprised upon learning that it was incorrect. Although this unproductively tenacious commitment on the part of subjects to a fallacious rule could have been readily avoided through the conduct of a single falsifying test (e.g., 6, 3, 1), subjects consistently sought evidence which would "confirm" their erroneous rules.

Mahoney and DeMonbreun (1977) borrowed this task from Wason in order to compare the critical reasoning skills of scientists with those of relatively uneducated Protestant ministers. Most of these subjects did not solve the problem and only two from among those who did were errorless (both ministers). Those subjects who did disconfirm were substantially more successful in discovering the rule. Neither scientists nor nonscientists, however, tended to employ disconfirmation. More than 85% of the triads generated were confirmatory. Scientists were indistinguishable from nonscientists; the frequencies of confirmatory experiments were not significantly different.

In another task designed by Wason (1968) known as the selection task or four card problem, subjects are asked to establish the truth value of a conditional proposition such as the following: "If a card has a vowel on one side, then it has an even number on the other side." Subjects must choose, from among a set of four cards, those cards, and only those cards, critical to this determination. Each of the four cards for this particular conditional has a letter on one side and a number on the other side. When the subject begins the task he is confronted with a vowel, a consonant, an even number, and an odd number. The essential features of the problem are reproduced in Table 1. The correct solution requires the selection of cards numbered "I" and "IV," that is, the vowel and the odd number. In order to appreciate the logic of this assertion, it is necessary to recognize that if a vowel and an odd number were to appear on the same card, then the hypothesis would be false; if they did not occur on the same card, the hypothesis would be true. The only cards which have the potential for revealing which set of circumstances actually holds are I and IV. Put more generally, the solution requires the insight that the odd number could falsify, and hence should be chosen, and that the even number could not falsify, and hence should not be chosen.

The typical subject selects the vowel and the even number, apparently failing to realize that the choice of the even number

Table 1

Analogue Hypothesis-Testing Problem

Assume that the four boxes which are presented below are actually cards which each have a letter on one side and a number on the other side. You are asked to test the hypothesis that--for these four cards--if a vowel appears on one side, then an even number will appear on the other side. Your "testing," of course, will involve turning one or more cards over.

I	II	III	IV
e	m	8	7

Are any of the cards irrelevant to the hypothesis?

(Circle) I II III IV No

Are any of the cards critical to the hypothesis?

(Circle) I II III IV No

Which cards would you turn over to test the hypothesis?

(Circle) I II III IV No

Note. From "From Ethics to Logic: A Survey of Scientists" by M.J. Mahoney and T.P. Kimper, Scientist as Subject: The Psychological Imperative. Cambridge, Mass.: Ballinger, 1976, p. 189.

merely provides irrelevant confirmation of the conditional. Simultaneously, the subject fails to select the odd number, as indicated above, a card with the potential to falsify the test sentence. Thus, subjects seem to be manifesting a bias toward confirmation. Wason and Golding (1974) have subsequently demonstrated that errors on this task are not dependent on test sentences with the surface structure of conditionals.

More recent research (Evans, 1972; Evans & Lynch, 1973; Wason & Evans, 1975) suggests that errors on the four card selection task are a result not of confirmatory bias but of a tendency to "match" the cards mentioned in the test sentence with the actual cards. Evans and Lynch (1973) have demonstrated that when subjects are presented with a conditional in which the consequent has been negated they tend to solve the problem correctly. That is, when subjects are presented with the test sentence, "If the card has a vowel on one side, then it does not have an odd number on the other side," they tend to solve the problem correctly. But the effect of a negative was found to be the same when the antecedent rather than the consequent was negated. That is, when the test sentence read, "If the card does not have a consonant on one side, then it has an even number on the other side," subjects selected the consonant and the even number, which is logically inappropriate. It appears that subjects select whatever cards are mentioned in the test sentence, at least when the material is presented in abstract terms (Johnson-Laird, Legrenzi, & Legrenzi, 1972). Errors may thus reflect "matching," rather than confirmatory, bias.

The debate regarding the diagnostic validity of this task with respect to confirmatory bias is not entirely resolved. Johnson-Laird and Wason (1970) define insight into the problem in accordance with the recognition that falsification of the conditional is more relevant than is verification. Goodwin and Wason (1972; cited in Wason & Evans, 1975) found a correlation between postulated degree of insight (as assessed by card choice) and the tendency to cite falsification as the reason for one's choice.

Confirmation Bias in a Simulated Research Environment. Mynatt, Doherty, and Tweney (1977) investigated problem solving in a setting designed to resemble the conditions under which actual science is conducted. Through the use of an on-line computer, a complex, dynamic environment was designed which permitted participants to generate tests of hypotheses. Subjects, who were undergraduate college students, selected their own experiments from among a larger pool, some of which could produce only confirmatory evidence for the hypothesis while others allowed the subject to test alternative hypotheses. The display presented stationary figures of three shapes (triangles, squares, or discs), with two brightness levels. A programmed keyboard command allowed subjects to move a small lighted dot or "particle" across the screen toward any point on the screen. A circular, non-visible boundary extended beyond the geometric centers of only low brightness figures. Whenever a particle encountered a boundary, its motion ceased. No other aspect of the stationary figures, such as size, shape, location, etc., affected particle motion. Hence, these were irrelevant cues. Mynatt et al. found strong evidence for a confirmation bias; subjects failed to choose environments which allowed for tests of hypotheses alternative to those initially selected by subjects. (Due to the relative salience of "shape" over size, location, and brightness cues most subjects generated hypotheses such as, "triangles deflect particles," and proceeded to test these hypotheses in a strictly confirmatory manner. Since numerous triangles could be found which were of low brightness, subjects had no difficulty "verifying" their erroneous hypotheses.)

In a subsequent study (Mynatt, Doherty, & Tweney, 1978a) subjects received extensive instruction in hypothesis falsification strategies (Popper, 1959, 1962) and in Platt's (1964) "strong inference" strategy prior to their interaction with the computer-simulated environment. Even though subjects were upper division undergraduates majoring in science (the mean number of science and mathematics courses taken was 19.3) and half had been instructed in disconfirmation, confirmatory strategies were overwhelmingly the rule. Subjects either abandoned a disconfirmed hypothesis but returned to it later, revised it in an attempt to account for the anomalous data, or ignored the disconfirmation and went on testing the same hypothesis.

Research with Professional Scientists. With the exception of the Mahoney et al. (1977) study which used professional scientists, all of the investigations described so far have employed undergraduate students as subjects.

Einhorn and Hogarth (1978) examined the inference strategies of persons known to have received extensive training in the logic of falsification. Twenty-three statisticians (faculty and graduate students in the department of statistics at the University of London) were required to solve a problem which concerned checking the claim, made by a consultant, of accurate predictive ability regarding rises and falls in a particular market. The problem is reproduced in Table 2. The correct answer is to select Responses 1 and 4, as only these responses permit disconfirmation of the consultant's claim. Twelve of the 23 participants requested a single piece of confirmatory data (Response 1 or 3) and only 5 people responded correctly to the problem, requesting both data necessary for falsification. Thus, it appears that the tendency to confine one's search to evidence consistent with the hypothesis under test is not limited to the judgments of laypersons.

Related Research

Several recent authors (Kern, 1981; Mynatt, Doherty, & Tweney, 1978b) have attempted to increase parsimony in the judgment and decision-making literature by suggesting that certain apparently disparate biases in intuitive inference might actually be viewed as varied manifestations of a single cognitive tendency toward confirmation of hypothesized relations. An extensive reinterpretation of the judgment literature in light of empirical research on confirmatory bias is contained elsewhere (Kern, 1981). For purposes of the present investigation it is sufficient to note that understanding of the following judgment heuristics might be increased by the view that each (to some extent) represents a manifestation of confirmation bias in the context of a different inference task: (1) the illusory correlation effect (Chapman, 1967; Chapman & Chapman, 1967, 1969; Golding &

Table 2

Stock Market Problem

It is claimed that when a particular consultant says the market will rise (i.e., a favorable report), it always does rise. You are required to check the consultant's claim and can observe any of the outcomes or predictions associated with the following:

1. favorable report.
2. unfavorable report.
3. rise in the market.
4. fall in the market.

What is the minimum evidence you would need to check the consultant's claim? Respond by circling the appropriate statement number(s).

Note. From "Confidence in Judgment: Persistence in the Illusion of Validity" by H.J. Einhorn and R.M. Hogarth, Psychological Review, 1978, 85, pp. 399-400.

Rorer, 1962; Smedslund, 1963; Ward & Jenkens, 1965); (2) illusory correlation in behavioral trait assessments (D'Andrade, 1965, 1974; Newcomb, 1931; Shweder, 1977); (3) implicit personality theory (Jackson, 1968; Koltuv, 1962; Mirels, 1976; Mischel, 1968, 1973, 1977; Norman & Goldberg, 1966; Passini & Norman, 1966; Snyder, Tanke, & Berscheid, 1977); (4) social stereotyping (Hamilton & Gifford, 1976; Snyder et al., 1977); (5) the pseudodiagnosticity effect (Doherty, Mynatt, Tweney, & Schiavo, 1979; Kern & Doherty, 1981); (6) belief in the law of small numbers (Tversky & Kahnemann, 1971); (7) the primacy effect (Anderson & Norman, 1964; Asch, 1946; Kanouse, 1972; cited in Fischoff, 1976; Peterson & DuCharme, 1967; Pruitt, 1961); (8) the inertia effect (Pitz, Downing, & Reinhold, 1967); (9) retrospective memory distortions (Bartlett, 1932; Bransford & Franks, 1971; Cofer, 1973; Franks & Bransford, 1971; Hastorf & Cantril, 1954; Loftus, 1975; Loftus & Palmer, 1974; Snyder & Uranowitz, 1978; Tzeng, 1972; Zangwill, 1972); (10) hindsight bias (Fischoff & Beyth, 1975); (11) inflated likelihood estimates resulting from the construction of causal schemata (Ross, Lepper, Strack, & Steinmetz, 1977); (12) "freezing" of hypotheses in clinical diagnosis (Dailey, 1952; Meehl, 1960; cited in Wiggins, 1973; Sines, 1959); and (13) perseverance of self perceptions in the face of disconfirming information (Ross, Lepper, & Hubbard, 1975).

What are the implications of an extensive body of literature suggesting a cognitive bias toward confirmation of our conjectures? If individuals do selectively emphasize events consistent with already held beliefs, are they necessarily led astray by this bias? Despite an accumulating body of empirical evidence that people find falsification strategies difficult to employ, philosophers of science (Hume, 1955; Platt, 1964; Popper, 1959, 1962) have long argued that confirmatory strategies are insufficient for determination of the truth status of a scientific hypothesis. Because of the foundational role which such arguments have played in the development of prescriptive norms for the conduct of research, a brief consideration of philosophical analyses of empirical hypothesis testing is necessary to the presentation of a rationale for the present investigation. Understanding of the formal arguments asserting a greater explanatory force for disconfirmatory inference requires, in turn, familiarity with certain principles of deductive logic.

Reason in Research: Formal Arguments Regarding Propositional Logic in Empirical Science

The test implications of a hypothesis are normally stated in conditional terms. That is, the prediction is made that, under certain specified test conditions, an outcome of a certain type will occur. In the typical strategy of theory evaluation the scientist formulates some hypothesis (p) which predicts the occurrence or existence of some phenomenon (q). In the social and biological sciences, for example, if p is a statement about a population and q is a statement about a sample, then there exists a relationship "if p, then q" (Bakan, 1967). What is observed is whether q is true or false. Conditional propositions formulated by researchers in the physical sciences typically involve the prediction of functional relations between quantities, which are stated in terms of a function rule. A simple example would be "c = v/r" or "if voltage equals v and resistance equals r, then current will equal c."

According to the inferential rules of propositional logic, a conditional statement of the form "if p, then q" is false "only in the case in which the antecedent p is true and the consequent q is false" (Wason, 1964, p. 30). If the conditional statement is true, it follows that if p also is true, then q must be true. "Any argument of this form is valid and is said to be in the affirmative mood or modus ponens" (Copi, 1978, p. 251). However, given the same true conditional (if p, then q) then the inference that if q is true then p must be true, is fallacious. "Any argument of this form is said to commit the fallacy of affirming the consequent" (Copi, 1978, p. 251).

What might induce such fallacious reasoning? The scientist reasons (implicitly or explicitly) that "if my hypothesis (p) is true then I should observe q." When he subsequently observes q he commits the fallacy of affirming the consequent if he concludes that this observation "verifies" his hypothesis. While it is in fact the case that the observation of q renders p more credible, "there is no definitive conclusion: the verification of its consequence q does not prove the conjecture p" (Polya, 1954, p. 4).

In addition to modus ponens, formal models of deductive inference define a second valid form of conditional inference, modus tollens. In modus tollens, given the conditional statement "if p, then q," if q is false then p must be false. In contrast, the inference that if p is false, then q must be false, is fallacious. "Any argument of this form is said to commit the fallacy of denying the antecedent" (Copi, 1978, p. 252).

As noted, given a conditional proposition of the form "if p, then q," what is observed in empirical work is whether q is true or false (this determination is typically made in probabilistic terms in the social and biological sciences). This has led philosophers of science to conclude on formal grounds that disconfirmation (i.e., not-q, therefore not-p) is the only form of valid conclusive inference in theory and hypothesis testing (Platt, 1964; Popper, 1959). From this follows the assertion that although successful predictions "may spawn subjective confidence [in the] theory or hypothesis from which they were derived,...it is only unsuccessful predictions which have conclusive logical implications" (Mahoney, 1976, p. 139).

But recent empirical work suggests that scientists leading successful research careers in physics, the biological sciences, and psychology have a poor understanding of propositional logic (Kern, Mirels, & Hinshaw, 1980). What accounts for the research contributions of these individuals if not their appreciation of logical propositions asserted by philosophers to be fundamental to the progress of science? In a larger sense, what accounts for the remarkable success of scientific research as a whole if in fact poor understanding of propositional inference is as widespread as investigations of professional scientists would suggest (Einhorn & Hogarth, 1978; Mahoney & Kimper, 1976)?

Despite their logical force and internal consistency, the "rules of deduction" outlined above give only a rough indication of the actual relationship between a hypothesis and its test implications in empirical science. When an investigator undertakes the observational test of a scientific hypothesis (H), he can rarely derive his test implications (I) from H alone.[1] In addition to H, a further set of complex auxiliary assumptions or auxiliary hypotheses are required.

Auxiliary Assumptions in Hypothetico-Deductive Inference

Discussion of the role of auxiliary assumptions in the formulation of an empirical prediction can best be introduced by means of a concrete example. Consider an account (Hempel, 1966) of Semmelweis' formulation of the germ theory of disease. Between the years of 1844 and 1848 Ignaz Semmelweis was a physician in the maternity division of the Vienna General Hospital. Between 1844 and 1846 approximately 10% of the women who delivered their babies in this division died of puerperal or "childbed" fever. Semmelweis attempted to determine the cause of childbed fever by subjecting various conjectured explanations for the disease to specific empirical tests. After rejecting a series of unproductive hypotheses related to variables such as overcrowding, the appearance of a priest before the onset of illness, the position of the woman during delivery, etc., Semmelweis serendipitously hit upon the conjecture that childbed fever was caused by contamination with "infectious matter" of the patient by the doctor performing the delivery. He reasoned that childbed fever could thus be prevented by chemically destroying the infectious material adhering to the doctor's hands. This led to the test implication that if

[1]Note that the symbols "p" and "q," typically used by formal logicians in discussions of propositional inference, will hereafter be replaced by the symbols "H" and "I," corresponding to "hypothesis" and "test implications," respectively.

attending physicians were to wash their hands in chlorinated lime before delivery, then maternal mortality from childbed fever would be reduced. The mortality from puerperal fever decreased promptly following issuance of a medical order requiring physicians to wash their hands in chlorinated lime before patient examination. But note that Semmelweis' test implication was not deduced from his theoretical hypothesis alone; rather, its derivation was based on the additional premise that a chlorinated lime solution, unlike simple soap and water, had the capacity to destroy infectious matter. This premise played the role of an auxiliary assumption in his complete formulation.

It follows from application of modus tollens that if H alone implies I, and empirical test has shown I to be false, then H must also be false. But when I is derived from H in conjunction with one or more auxiliaries (A), then modus tollens reads as follows: "If (H·A) implies I and we have falsified the implication (I) observationally, we can conclude not-(H·A)." Obviously, this result does not falsify the substantive hypothesis but rather it falsifies the conjunction of the substantive hypothesis with an auxiliary. Thus, the failure to observe a predicted test implication only proves not-H, not-A, or both, but we don't know which. Note also that the number of auxiliaries which may be conjoined with a focal hypothesis in any empirical test is, in principle, infinite, so that failure to observe a predicted outcome may serve to prove not-H or not-A_1 or not-A_2 or not-A_3...or not-A_n. This implies that falsification will leave the researcher with an increasingly ambiguous picture of the phenomenon under investigation as the number of auxiliary hypotheses increases.

To emphasize the significance of auxiliary hypotheses for empirical testing let me return to the Semmelweis example. What if the antiseptic measure employed had not produced a decline in mortality? The focal hypothesis might still have been true; the failure of the prediction having resulted from inefficacy of chlorinated lime as an antiseptic. This argument need not be made contrary to historical fact. The astronomer Brahe rejected the heliocentric theory of planetary motion on the basis of a faulty auxiliary assumption. He reasoned that if the Copernican hypothesis were true, then an observer standing on earth and staring at a fixed star at a fixed time of day should see the position of the star gradually change. Because the telescope had not been invented when Brahe began his search for parallactic motions, even the most precise instruments available in his time were insufficiently sensitive to detect the predicted changes. Brahe felt compelled to reject the heliocentric theory, not because the test implication was wrong, but because he had made the faulty auxiliary assumption that fixed stars are sufficiently close to the earth's surface to yield parallactic movements detectable with his instrumentation.

This last example serves to illustrate another important point with respect to auxiliary hypotheses. Recall that in the introduction to this section it was asserted that the test implications of a hypothesis are typically stated in conditional terms. That is, the prediction is made that <u>under certain specified test conditions</u>, an outcome of a certain type will occur. These conditions are formally symbolized as "C." Suppose a test implication has been derived from

H and set A of auxiliary hypotheses. The experimenter's task is then defined as checking whether or not some predicted outcome occurs in a situation in which the conditions C are realized. But what if the specified conditions are not met? What if, for example, the instrumentation is faulty or insufficiently sensitive? The predicted outcome may fail to occur even though H and A are true. Thus, the total set of auxiliary assumptions includes the supposition that the test instrumentation meets the specified conditions C.

The Present Investigation: Rationale

It is the central thesis of this investigation that the presence of auxiliary assumptions during formulation of an empirical prediction decreases the likelihood of disconfirmatory reasoning on the part of the experimenter. That is, it is contended that the likelihood of confirmatory strategies in scientific hypothesis testing increases as a function of the number of auxiliary assumptions conjoined with a focal hypothesis in a given empirical test.

Evidence for this thesis must be derived from research examining other variables in the scientific hypothesis-testing process, since the number of auxiliaries has never been independently manipulated. One relevant study investigated the behavior of mathematicians engaged in testing a conjecture (Markowitz & Tweney, 1981). The study was provoked, in part, by Hadamard's (1945, cited in Markowitz et al., 1981, p. 2) discussion of the role of errors in mathematics. He asserted that while a good mathematician is quick to perceive and correct inferential errors, the experimental scientist is much less likely to recognize such errors. Markowitz undertook the first observational investigation of mathematical discovery. He wanted to assess what inferential methods mathematicians use when they investigate conjectures, more specifically, how mathematicians proceed after they find a counterexample (see Table 3) to a proposed conjecture. Do they abandon the conjecture? Modify it appropriately? Perseverate without revision?

Graduate students and faculty in university departments of applied statistics and mathematics spent two hours investigating a conjecture in number theory. Relevant portions of the problem are reproduced in Table 3. As they worked, subjects were asked to verbalize their reasoning at every step of the problem. From these responses and written materials produced by subjects, a protocol of the individual's inference strategies was constructed. A clear outcome of this investigation was that subjects successfully employed disconfirmation to modify or reject conjectures. "In only one instance was a conjecture retested without revision after a counterexample was found" (Markowitz et al., 1981, p. 5). This finding lends support to Hadamard's (1945) assertion that disconfirmation plays a different role in mathematical versus scientific inference. Recall that Mynatt et al. (1978a) found that disconfirmation rarely led subjects in the computer simulated research environment to permanently abandon a hypothesis. To repeat, subjects either abandoned a disconfirmed hypothesis but returned to it later, revised it in an attempt to account for the anomalous data, or ignored the disconfirmation and went on testing the same hypothesis. In a related vein, Mitroff (1974), by means of extensive interviews with 40 NASA scientists,

Table 3

The Abundant Number Conjecture

<u>Definition:</u>	An abundant number is a number whose positive divisors sum to more than twice the value of the number.				
<u>Examples:</u>	<u>Divisors</u>	<u>Sum of Divisors</u>			
12	1,2,3,4,6,12	28	28		2 x 12
18	1,2,3,6,9,18	39	39		2 x 18
24	1,2,3,4,6,8,12,24	60	60		2 x 24
<u>Nonexamples:</u>					
4	1,2,4	7	7		2 x 4
8	1,2,4,8	15	15		2 x 8
10	1,2,5,10	18	18		2 x 10
<u>Conjecture:</u>	A number is abundant if and only if it is a multiple of 6.				
<u>Subconjecture 1:</u>	If a number is a multiple of 6, it is abundant.				
<u>Counterexample</u>	<u>Divisors</u>	<u>Sum of Divisors</u>			
6	1,2,3,6	12	12		2 x 6
(only counterexample that exists)					
<u>Subconjecture 2:</u>	If a number is abundant, it is divisible by 6.				
<u>Counterexamples</u>	<u>Divisors</u>	<u>Sum of Divisors</u>			
20	1,2,4,5,10,20	42	42		2 x 20
40	1,2,4,5,8,10,20,40	90	90		2 x 40
56	1,2,4,7,8,14,28,56	120	120		2 x 56
(infinitely many counterexamples exist)					

found that commitment to confirming one's own theoretical position was viewed by many of these researchers as desirable and necessary. They argued that without such commitment, many relevant, original, but not fully developed ideas would be abandoned due to premature falsification. Markowitz et al. assert that the differences in inferential strategy employed by mathematicians and empirical scientists may reflect differences in the ambiguity of the relationship between data and hypothesis in the two disciplines.

How does a mathematical counterexample differ from a failed empirical prediction? A single counterexample to a mathematical conjecture conclusively refutes the conjecture. Note Subconjecture 2 from Table 1: "If a number is abundant, it is divisible by 6." The number 20 is abundant since its positive divisors sum to 42. But it is a counterexample to Subconjecture 2 since it is not divisible by 6. Likewise, it may seem initially that when the test implication of a scientific hypothesis is formulated in terms of a function rule (e.g., "c = v/r"), then a single deviant observation serves to refute the hypothesis. For example, if the known ratio of voltage to resistance does not equal the measured value for current on a single occasion then it would seem logical to assume that we have conclusive refutation of our hypothesized relation. There appears to be no ambiguity in the relation of data to hypothesis in either our mathematical or our scientific example. But note that in the c = v/r case, a deviant value for c could reflect a failure in the instrumentation used to measure current. Thus, the experimenter might say to himself, "Perhaps it is something about this current reading in particular which is anomalous. If I take another reading using a different instrument I may get the value originally predicted." Such ambiguity would never arise for the mathematician (at least in most areas of mathematics). An individual testing a conjecture would never be led to assert, "Perhaps it is something about <u>this 20</u> in particular..." or "Perhaps if I used <u>another</u> form of addition..." Unless the investigator had made an arithmetic error (a possibility which can always be checked conclusively) a single counterexample is sufficient.

Thus far we have compared mathematical with empirical science only in the instance in which the empirical prediction is stated in terms of a function rule. Consider the prediction of the social scientist that an inverse correlation exists between "compulsivity" and susceptibility to post-hypnotic suggestion. The investigator assigns high and low compulsives (as assessed by scale 7 of the MMPI) to separate groups and then measures the number of unusual acts (out of a total of five) performed by these subjects following hypnotic suggestion. If the investigator fails to observe the predicted association, he has refuted the conjunction of his substantive theory, T, with the following auxiliary assumptions, to name just a few: (1) There exist enduring, cross-situational consistencies in the behavior of individuals ("traits"); (2) These consistencies in behavior can be reliably and validly assessed by means of the MMPI; (3) The only variable on which the two experimental groups differed prior to hypnotic induction was in degree of "compulsivity"; (4) The hypnosis manipulation was successful; (5) Data collection by research assistants was entirely reliable; etc. From a commonsense, nonformal standpoint this problem introduces a "quantum leap" in the ambiguity of the relationship between data and hypothesis when compared with the mathematical and scientific examples depicted above.

What variables account for this difference? First, the sheer number of auxiliaries is much greater in the latter instance than in the former two. A second difference relates to a difficulty confronting the social scientist who attempts to state, before the fact, what would constitute strong falsification of a theory. Meehl (1978) has made the astute observation that in the physical sciences there is an intimate connection between a theory under test and the auxiliary assumptions simultaneously subjected to falsification via modus tollens. This "intimate" connection is necessitated in the physical sciences by an overlap in the inner components of the propositions. The nature of this conceptual overlap is epitomized by means of an example:

> There is a complicated, well-developed, and highly corroborated theory of how a cyclotron works, and the subject matter of that auxiliary 'theory of the instrument' is for the most part identical to the subject matter of the physical theories concerning nuclear particles, and so on, being investigated by the physicist. Devices for bringing about a state of affairs, for isolating the system under study, and for observing what occurs as a result are all themselves legitimated by theory. (Meehl, 1978, p. 819)

Unfortunately, no such propositional overlap exists in "soft" psychology. Typically, the auxiliary assumptions (such as the assumption that the MMPI is a valid indicator of a "trait" labeled compulsivity) are in no way necessitated by the substantive theory under test. Thus, the composition of the auxiliaries becomes somewhat arbitrary and would likely be considered "inelegant" or "unaesthetic" using the evaluation criteria of a physical scientist. Additionally, Sarbin, Taft, and Bailey (1960) assert that evaluation of the truth status of an auxiliary assumption is made more difficult in the social sciences by the fact that terms frequently have multiple referents and provide the basis for chained associations of meaning.

Thus it would appear that the nature of the relationship between the theory under test and the auxiliary assumptions simultaneously subjected to empirical evaluation is a complex one, varying along a number of dimensions, all of which point to the existence of a continuum ranging from the mathematical, to the physical, to the social, sciences.

This has led a few recent authors (Greenwald, 1981; Meehl, 1978) to assert that the use of disconfirmatory inference strategies may be differentially appropriate in the physical versus the social sciences. It is Meehl's contention that, given the weak connection between the theoretical and auxiliary components of propositions formulated by many social scientists, it would be a <u>very strange coincidence</u> if a substantive theory lacking verisimilitude could be conjoined with arbitrary auxiliary assumptions to yield results consistent with a precise empirical prediction. Meehl exhorts social scientists to attempt the prediction of numeric point values, or, in the event that one's substantive theory is in too primitive a stage of development for this, then the prediction of ranges, orderings of dependent variable values, and the specification of function forms is recommended. Note that Meehl does not assert that all forms of confirmatory inference are equally powerful in psychological research. Rather he believes that while confirmation via the prediction of a directional difference between means represents a weak corroboration of a psychological theory, confirmation via estimation of numeric point values has strong corroborative power. The essential point is that Meehl believes there are important differences between the exact and the inexact sciences with respect to how and when confirmatory versus disconfirmatory methodology should be applied. Perhaps, when used appropriately, confirmatory inference represents an extremely powerful methodology.

With respect to the present investigation, Meehl's speculations suggest that claims by philosophers of science that disconfirmation is somehow universally "good" while confirmation is universally "bad," may be simplistic. Perhaps there are elements of the task situation which dictate the appropriateness of a confirmatory versus a disconfirmatory inferential approach. It may represent a step forward in understanding the literature on confirmatory bias to view this phenomenon as one more useful judgmental heuristic which is, under a number of circumstances, over-extended into problem domains in which falsification strategies would be more effective.

CHAPTER III

Design and Hypotheses

It was the purpose of the present investigation to examine the impact of auxiliary assumptions on data selection strategies in scientific hypothesis testing. It was predicted that, as the number of auxiliary assumptions conjoined with a focal hypothesis increased, the likelihood that investigator-subjects would employ confirmatory inference strategies would increase correspondingly. Conversely, it was predicted that falsification strategies would be utilized most frequently and most effectively in situations where no auxiliary hypotheses existed.

It was indicated above (p. 6) that confirmatory bias may manifest itself in any one of four ways. For the purpose of incorporating a dependent variable into the present investigation which could be assessed with certainty and precision, the tendency to employ a verification strategy was operationally defined in terms of "inferential perseveration"; that is, the failure to abandon a favored hypothesis once it had been falsified.

There are two ways in which an investigator can perseverate with a favored, but empirically refuted, hypothesis. First, the individual can simply ignore the disconfirmation and go on testing the same hypothesis. Alternatively, in response to incoming disconfirmatory data, an investigator can elect to examine the truth status of an auxiliary assumption as a means of postponing examination of the truth status of the focal hypothesis. That is, whenever an empirical observation is made, the investigator asks, "Is this a theoretically meaningful result or is it simply an artifactual one related, perhaps, to a defect in my experimental methodology?" If an individual elects to examine the truth status of an auxiliary only in the face of disconfirmatory empirical results, while auxiliaries remain unquestioned in the face of confirmatory data, then the individual is manifesting one form of "inferential perseveration."

Mahoney (1977), in an experimental study of the peer review system, asked seventy-five journal reviewers to referee manuscripts which contained identical introductory and methodology sections, but which reported data which were either consistent or inconsistent with the reviewer's presumed theoretical perspective. These opposite sets of data defined "positive" and "negative results manuscripts," respectively. With identical experimental procedures, a manuscript with positive results was rated as methodologically superior to one reporting negative results. This supports the contention that the truth status of an auxiliary is more likely to be questioned when data are disconfirmatory, rather than confirmatory. Although Mahoney interprets this finding in motivational terms, one can evaluate this outcome as one manifestation of the difficulty of interpreting a disconfirmatory result in the context of ambiguity created by the presence of auxiliary assumptions.

In order to examine the impact of auxiliary assumptions on hypothesis testing it was necessary to isolate this variable from phenomena with which it tends to covary in the actual practice of science, such as the level of conceptual overlap in the propositions' inner components (see p. 22 above), or the nature of the scientific domain in which an empirical question is being asked (exact versus inexact science). An effective manipulation of auxiliary hypotheses required the separation of this variable from all others within the task domain.

In the present investigation, subjects were asked to test a concrete hypothesis in a setting modeled after a real research setting. The number and nature of auxiliary hypotheses (conjoined with the focal hypothesis) constituted independent variables. As indicated above, inferential perseveration can manifest itself as (1) failure to abandon a disconfirmed hypothesis or (2) a greater tendency to examine the truth status of an auxiliary assumption in the face of disconfirmatory, rather than confirmatory, data. Each of these behaviors constituted a separate dependent variable.

Note that the various phenomena which legitimately fall under the definitional umbrella term "auxiliary hypothesis" are quite diverse. Contrast, for example, Semmelweis' belief that chlorinated lime is an effective antiseptic with the belief that MMPI responses are a valid indicator of "compulsivity." Each of these hypotheses may, in turn, be contrasted with assumptions regarding the precision or level of functioning of instruments used in the manipulation and assessment of independent and/or dependent variables. Auxiliary assumptions of the former type, which relate to theoretical presuppositions, are not readily manipulated artificially because they may contradict (or unite with) numerous a priori beliefs which sophisticated subjects bring to the task. For example, graduate students in the natural sciences have certain beliefs about laws governing the physical universe which it would be "unnatural" for them to suspend for the purpose of participating in a psychological experiment. Note, however, that auxiliary assumptions of this type have an "all-or-none" effect on the data which is obtained by the investigator. Returning to the Semmelweis example, if chlorinated lime had not been an effective antiseptic, then the incidence of childbed fever would have been <u>totally</u> unaffected by the experimental manipulation employed. When auxiliary assumptions relating to theoretical presuppositions turn out to be false, the anticipated experimental outcome fails to occur, in an all-or-none, unambiguous fashion. In common-sense terms, the whole experiment is thrown "out of whack." Perhaps this effect can be replicated in an artificial environment by introducing instrumentation which functions in a totally unpredictable fashion on some proportion of the trials in which it is employed. This may capture the essence of what occurs on a concrete level in an experimental situation when an accurate focal hypothesis is conjoined with a false auxiliary assumption, but note that this manipulation does not contradict any a priori beliefs held by sophisticated subjects regarding physical reality. This all-or-none error must be distinguished from that which arises when some amount of random and/or systematic error is introduced into <u>each experimental datum</u> which is returned to the investigator. When some amount of irreducible error is contained in every observation received, then there is a violation of the auxiliary assumption that the test arrangement satisfies the specified test conditions C (Hempel, 1966).

Both types of such instrumentation error were manipulated in the present study in order to examine the impact of auxiliary assumptions on inference strategy. That is, two dimensions of data ambiguity were introduced which involved (1) some specifiable but irreducible random error in each observation received or (2) a complete breakdown of the instrumentation on a specified proportion of trials. The subject's task involved acquisition of empirical information about the lawful properties of an unexplored planet through several "layers" of data gathering equipment. It was hoped that this layering of instrumentation would make plausible to subjects the possibility of equipment failure or imprecision since, in actual science, it is generally true that the greater the physical separation between an experimenter and the phenomenon investigated, the less direct, immediate, and, perhaps, accurate empirical observations become.

Each of the two forms of error were varied independently in a 2 x 2 factorial design such that subjects exposed to Condition 1 received perfectly "true," error-free data each time an experimental observation was made; subjects exposed to Condition 2 received data which incorporated some amount of irreducible error into <u>every</u> observation obtained; subjects exposed to Condition 3 received data by means of an instrument which broke down on some <u>proportion</u> of all experimental trials which were undertaken; and subjects exposed to

Condition 4 received data which incorporated both forms of error just described--that is, some irreducible error was present in each observation and total instrumentation failure occurred on some proportion of all experimental trials. It was generally predicted that the introduction of both forms of error, singly and in combination, would yield an increase in the degree of inferential perseveration displayed by subjects. There was no basis for predictions regarding the magnitude of this increase or the possible interactive effects of combining the two forms of error in the fourth experimental condition. The general prediction regarding inferential perseveration followed from the premise that, when instrumentation error was introduced, the meaning of a failed prediction would become increasingly ambiguous, thus weakening the "refutation power" of a disconfirmatory result.

The design of the present study, in addition to providing a means for examining data evaluation strategy as a function of auxiliary hypotheses, also permitted examination of the impact of different forms of data error on inference. Because the empirical study of scientific hypothesis testing is a new domain in psychological research, very little is known about the impact of data error on reasoning strategies. It was conceivable that error which was present in every observation would have a different impact on subjects' evaluation of data than would error which was present on some trials but absent on others. No specific predictions in this regard were made.

In sum, two auxiliary hypotheses relating to the accuracy of experimental instrumentation were independently manipulated in a 2 x 2 design. The possible role of auxiliary assumptions in promoting a confirmatory inference strategy was assessed by evaluating the degree of "inferential perseveration" manifested by subjects in each experimental condition. Inferential perseveration was operationally defined as (1) refusal to abandon a hypothesis in the face of a failed empirical prediction and (2) a greater tendency to question the accuracy of instrumentation in the face of disconfirmatory data.

CHAPTER IV

Method

Subjects

Forty graduate students from the departments of physics (n = 16), chemistry (n = 12), and biology (n = 12) participated in the study. Mean number of years in graduate training at the time of participation was 3.8 (range = 1-6, including 3 post-doctoral fellows). Each student was paid $10.00 for participation.

Recruitment

After obtaining permission from their respective department chairmen and classroom instructors, the investigator solicited subjects during the first few minutes of graduate seminar meetings. She explained that the study was an empirical investigation of the scientific inference process, and that it involved computer simulation of a scientific hypothesis-testing problem. She indicated that no specialized knowledge was required for understanding or solution of the problem. Following this brief description of the task, a $10.00 inducement was offered to encourage participation. In all, 42 individuals were approached to obtain a sample of 40 participants.

Procedure and Design

As indicated in Chapter III, the independent manipulation of measurement and stochastic error yielded a total of four experimental conditions: (1) no error; (2) measurement error only; (3) stochastic error only; and (4) both measurement and stochastic error. An equal number of subjects within each discipline were assigned to each experimental cell by means of a randomized-blocks design procedure. (All four groups consisted of three biology students, three chemistry students, and four physics students.)

Instructions presented to participants in Condition 4 (measurement and stochastic error) are presented below in their entirety. Instructions to subjects in the other three conditions are completely derivable from these instructions, simply by deletion of portions relevant to one or both types of data error. (The full instructions for these conditions are presented in Appendix A.)

After being seated in front of the color monitor and keyboard of an Apple II minicomputer, subjects assigned to Condition 4 received the following instructions:

> You are a scientist, investigating an unexplored planet, Ethereus. Right now, you are orbiting Ethereus in a spaceship. From your spaceship, you can conduct a variety of controlled experiments. Previous research has shown that certain life forms exist on the planet, but the conditions which support these life forms are very poorly understood. Your research project will involve an attempt to determine what conditions promote the growth of a plant, the tribble, found in certain regions of Ethereus. The tribble was selected as the focus of this initial investigation because earlier work suggests that its survival depends only on the amount of moisture present in the soil. It is suspected that above a certain moisture content, tribbles grow. Below this moisture content, the plants die. Your task is to determine what this critical level is by systematically planting tribbles at various points on the planet's surface and

> seeing whether or not they survive at these locations. Each of the points you select for planting will correspond to a certain moisture level. The site of the investigation will be a 250,000 square-mile area encompassing a large portion of the planet's southern hemisphere. There is a display of this area on the computer console in front of you.

Subjects were shown the display illustrated in Figure 1.

> Fortunately, research has established that the distribution of moisture in the planet's soil is remarkably regular. The percentage of soil moisture on your viewing screen increases uniformly from west (left side of screen) to east (right side of screen). Preliminary data indicates that tribbles can survive only when the percentage of moisture in the soil equals or exceeds a certain level indicated by this dotted line (E indicated--see Figure 2). Your job is to conduct a more thorough investigation of the tentative hypothesis that tribbles can survive only at moisture levels equal to or exceeding the level displayed on the screen by planting them east or west of this line and observing whether they live or die.
>
> Each time you're ready to plant a tribble, a short horizontal line will appear on the screen. This line (E indicated to S--see Figure 3). represents an area in which you can plant your tribble. This area can be moved to the east or west of where it first appears very easily. If you want to move the area to the east, just hold down the "repeat" (REPT) key and press the arrow beneath it on the right (→). To move the area in a western direction, hold down the repeat key and press the arrow on the left (←).
>
> Why is the planting area represented by a horizontal line instead of a dot? Because you are orbiting the planet from a distance of 500 miles, you are not able to plant tribbles at a precise location. Rather, tribbles will land within 15 miles of the location at which you aim. Although the tribble may land somewhat east, west, north, or south of the location you specify, only

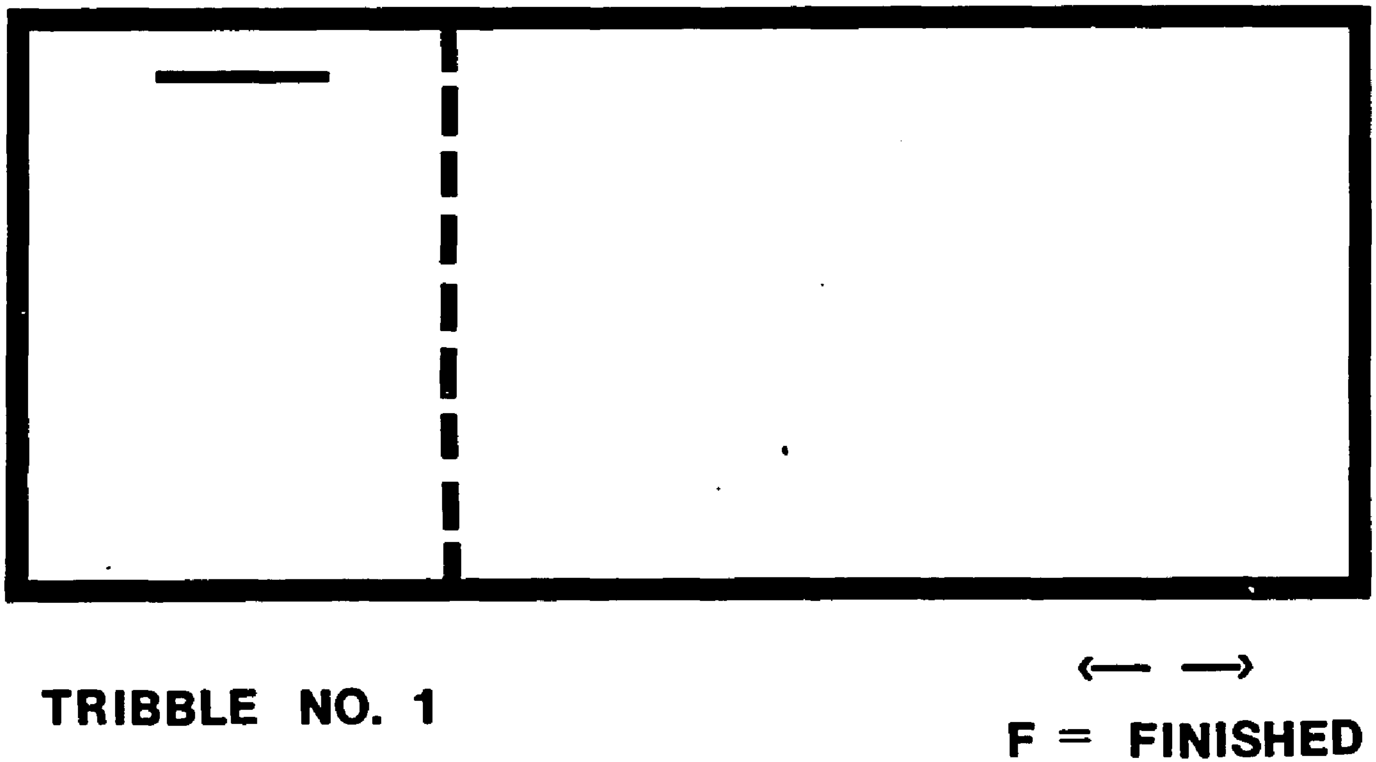


Figure 1: Illustration of initial display screen. The investigator indicated to S that the large rectangular area represented the site of the investigation.

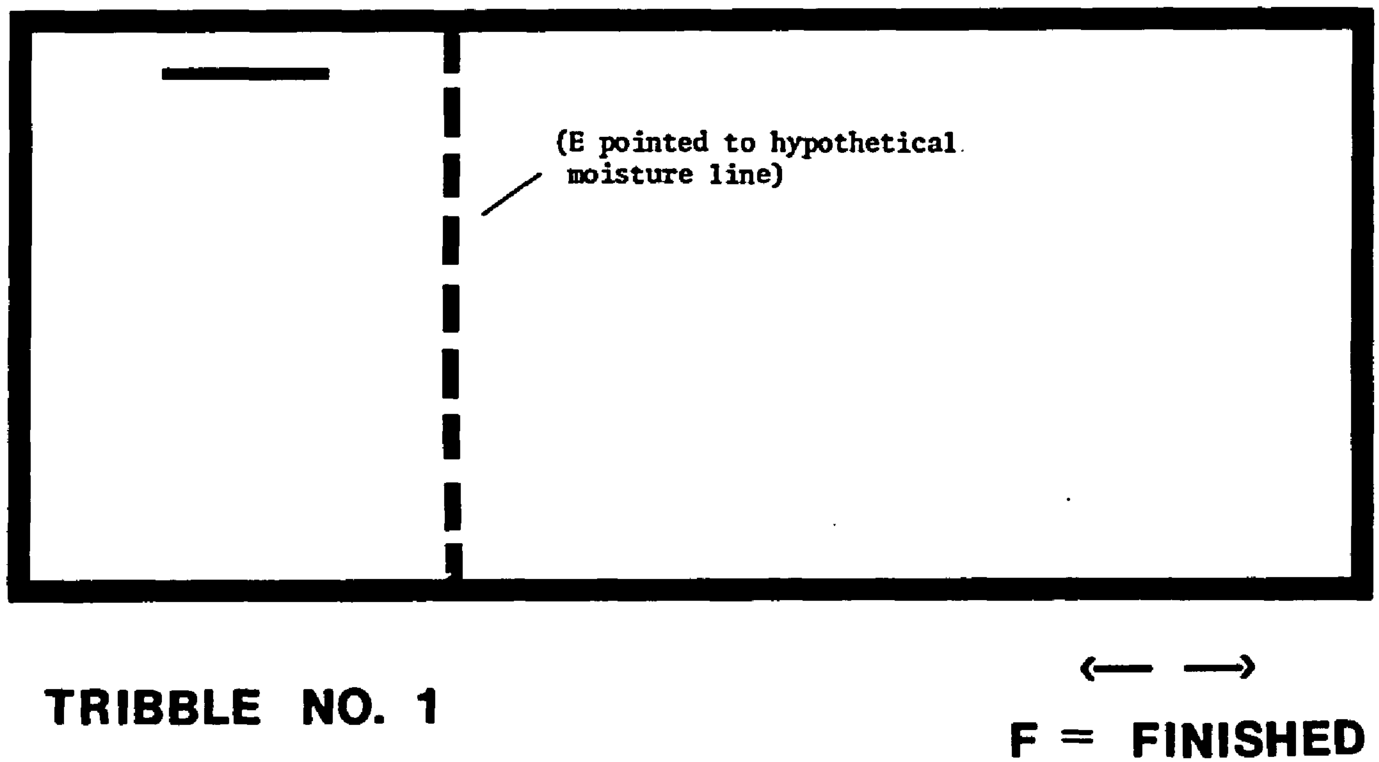


Figure 2: The investigator indicated the initial location of the hypothetical moisture line by pointing to it on the screen.

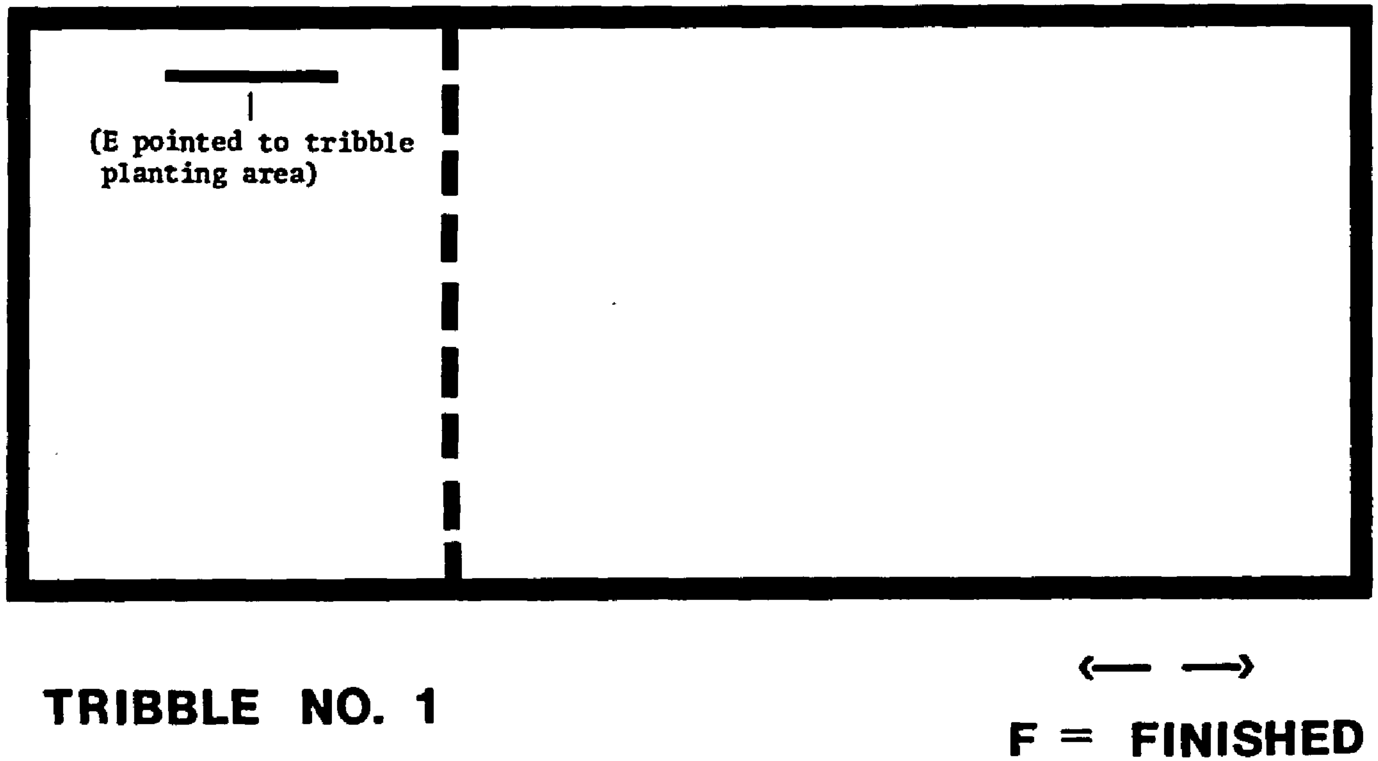


Figure 3: The investigator identified the tribble planting area by pointing.

> east-west error is of interest with respect to your moisture hypothesis, so this is the error you see represented on the screen by the horizontal line. By moving the line east or west on each trial you can specify a planting location. When the tribble is released, you will know that it has landed somewhere within the east-west range represented by the horizontal line, but you can be no more specific in your observations than this.

At this point, E demonstrated how to plant a tribble. When the tribble landed, E said, "There, the line turned green/purple, so that one lived/died. Does that make sense?"--See Figure 4 for an illustration the screen seen by S at this point.

> Also because you are orbiting the planet from a considerable distance, there are several layers of instrumentation involved in collecting data regarding whether or not your tribble survived. Specifically, your spaceship is equipped with a series of probes, one of which is sent down with every tribble you plant. Occasionally, the probe malfunctions and sends back random information totally unrelated to the true condition of the tribble. That is, what the probe says about the tribble bears no relation to its true condition (E explained). Previous work involving such probes indicates that the rate of malfunction is approximately 25%, or one in four. Because this is an average, however, your own rate of probe malfunction may be slightly greater or less than 25%. Fortunately, your spaceship is equipped with a device which can check on the functioning of the probe and tell you when you have received random data resulting from instrument malfunction. Because of the financial costs associated with probe checks, you will be allowed to employ it on only two occasions.
>
> You have resources available to plant a total of 8 tribbles.
>
> In order that I can better understand what you're thinking about as you perform the task,

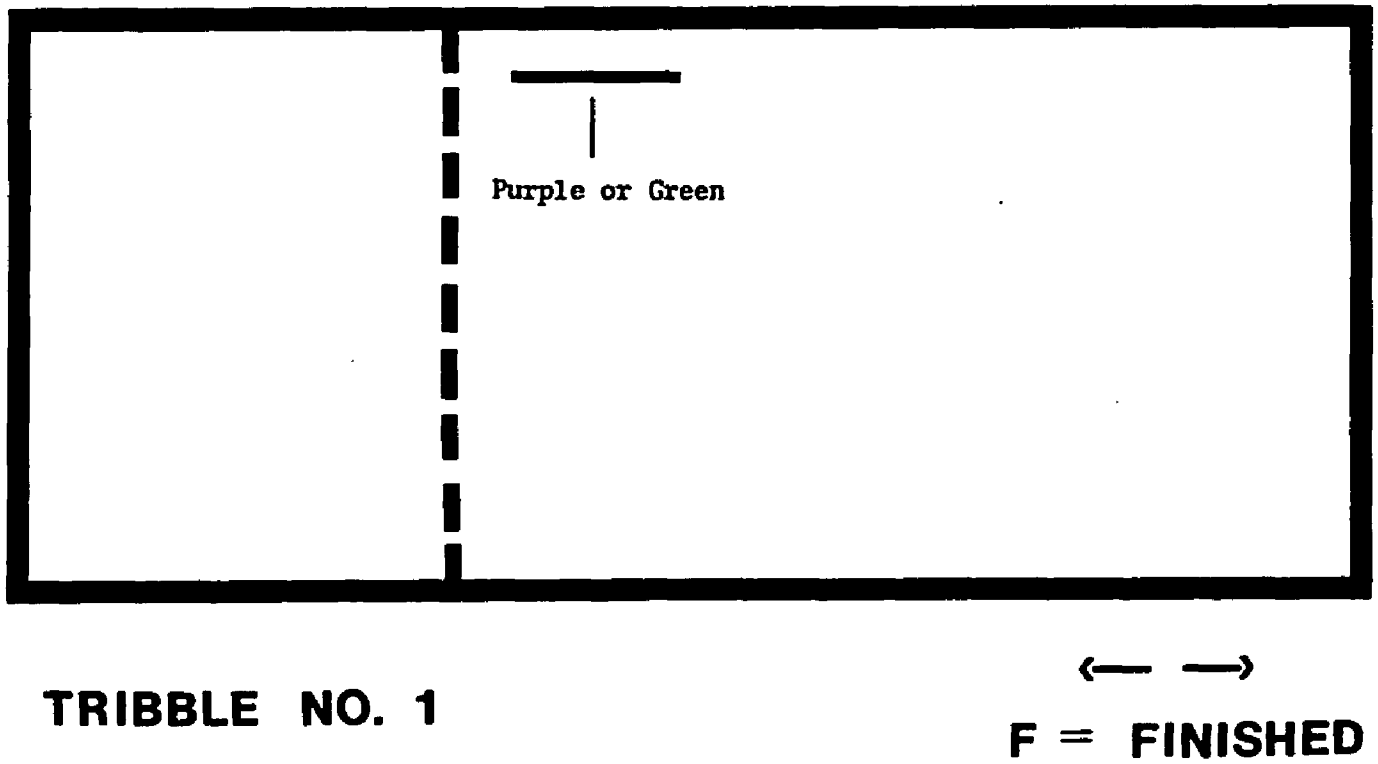


Figure 4: The investigator demonstrated how to plant a tribble, stating that a green horizontal line indicated a live tribble within the planting area, while a purple horizontal line indicated that the area contained a dead tribble.

I'd like you to go through the following steps each time you plant a tribble:

1. When the line appears, position it over the location at which you would like to plant. After the planting area has been selected, press the "F" key for FINISHED. (E demonstrated)

2. At the bottom of the screen it now says, "PREDICT RESULTS; D = DIE, L = LIVE." Based on the data you have available to you, I'd like you to try and predict whether the tribble you're about to plant will live or die. Initially, you may feel you don't have enough information to warrant a reasonable prediction, but I'd like you to do your best.

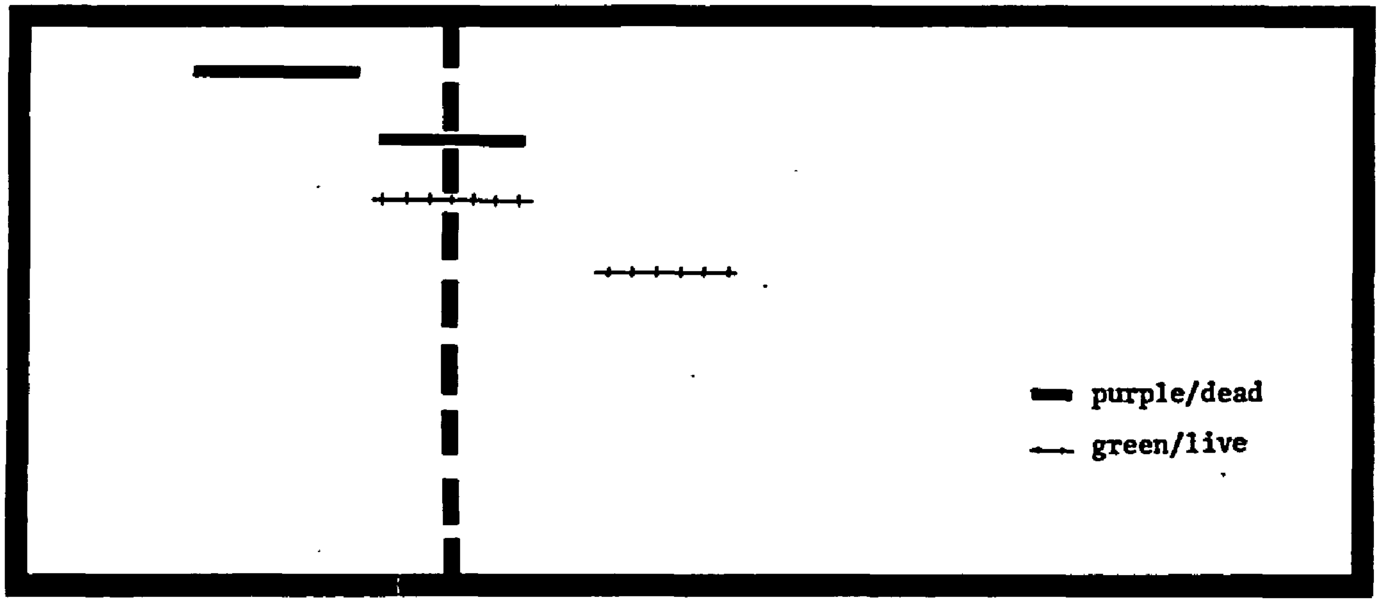
3. As soon as you've formulated your prediction, I'd like you to tell me, as best you can, the reason for your prediction.

4. Release the tribble by pressing the "L" or the "D" key. The tribble will then descend to the surface of the planet and either grow or die. The result will be radioed back to your space craft and displayed on your screen as either green or purple.

E demonstrated and then presented Figure 5 to subjects while making the inquiry displayed at the bottom of the Figure.

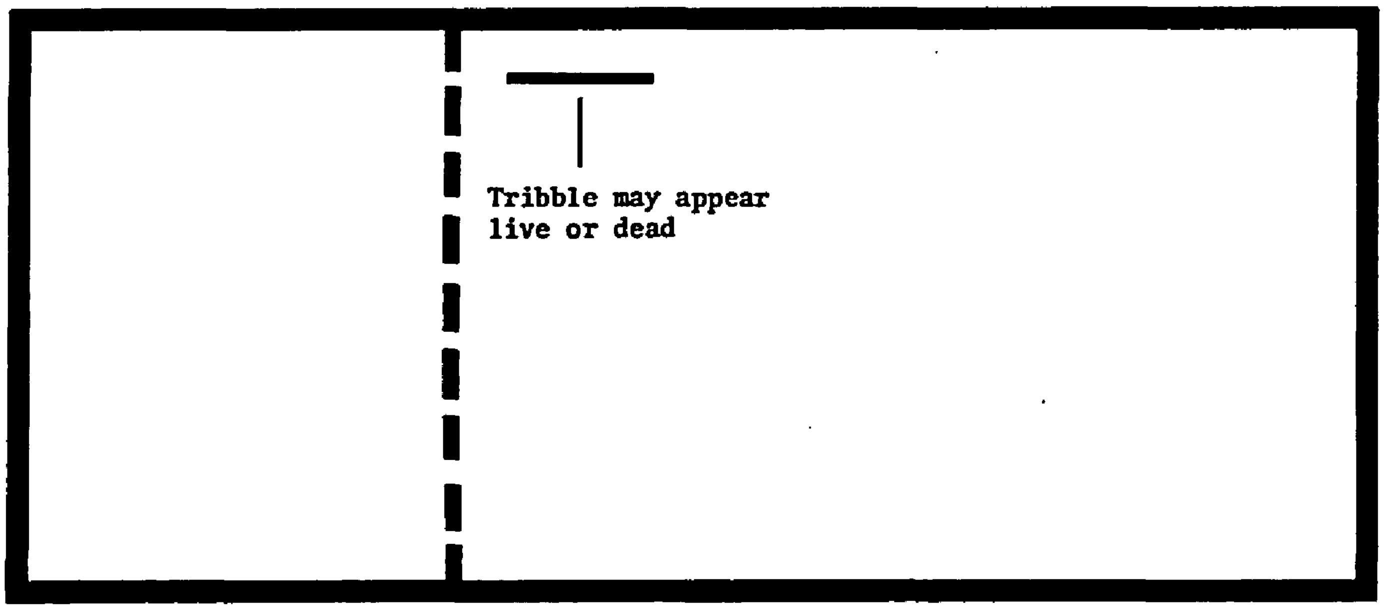
5. Next, decide whether or not you want to do a probe check on this trial. Indicate "Y" for YES or "N" for NO in accordance with your decision. (E demonstrated --see Figure 6) Remember, you have only two probe checks available.

After a decision had been made regarding the probe check, the display changed to that illustrated in Figure 7.



TRIBBLE NO. 4

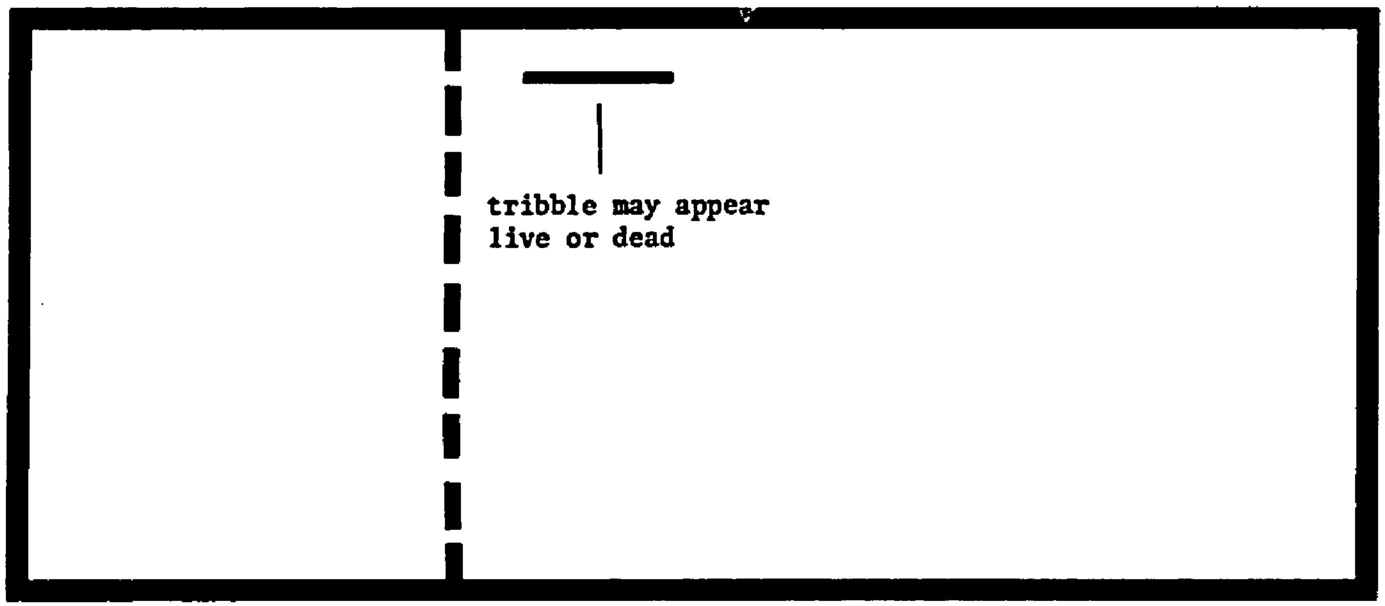
Figure 5: While showing the above Figure to Ss (with actual colors included), the investigator made the following inquiry: "Assuming the probe functioned properly on all four of the above trials, and assuming that the dotted line represents the true location of the critical moisture line, explain how the above data could have been generated." This question was designed to test Ss' understanding of the concept of measurement error.



TRIBBLE NO. 1

DO YOU WANT TO DO A PROBE CHECK ON THIS TRIAL? TYPE Y OR N

Figure 6: Illustration of the appearance of the display at the point of a possible probe check.



TRIBBLE NO. 1

DO YOU WANT TO CHANGE THE HYPOTHETICAL LINE? TYPE Y OR N

Figure 7: Illustration of the appearance of the display at the point of a possible hypothesis revision.

6. At this point, the lower portion of the computer screen says, "DO YOU WANT TO CHANGE THE HYPOTHETICAL LINE? TYPE Y OR N." What I'm interested in finding out is whether or not the data you've generated has led you to reject the location of the original critical moisture line in favor of a new location. I want you to relocate this line only when your data indicates that the line's present location is wrong and does not represent the actual critical moisture level. Do this by typing "Y" for YES, and repositioning the hypothesis line in the same manner as repositioning a tribble (by using the arrow keys and the repeat key). (E demonstrated--see Figure 7) When you've reached a point that represents your new working hypothesis about where the critical moisture line should be, type "F" for FINISHED and the computer will give you your next tribble. If you don't want to move the line, type "N" for NO and the computer will make your next tribble available for planting. This procedure will be followed for each tribble you plant until you've planted a total of 8 tribbles. Do you have any questions?

At this point, the investigator clarified areas on which the subject remained unclear.

> It would be helpful to me if you could "talk out loud" as you perform the task, sharing with me any inferences you may be making or reasoning strategies you may be using as you work. I'll be making a tape recording of any thoughts or perceptions you share, so that I can go back later and correlate these with specific responses you've made on the task.

Thus, the manipulation of measurement error (in Conditions 2 and 4) was accomplished through incorporation of landing error into each datum returned to S, while stochastic error was manipulated (in Conditions 3 and 4) through introduction of a probe which malfunctioned on

25% of experimental trials, returning randomly generated data on such trials. As noted above, instructions to subjects in Conditions 1 through 3, derived from those presented here through deletion of portions relevant to one or both types of data error, are contained in Appendix A. Briefly, subjects not exposed to measurement error were able to plant tribbles at a specific location, represented visually as a single dot on the display screen; subjects not exposed to stochastic error were never introduced to the concept of a malfunctioning probe, they were led to assume that data was obtained by direct observation.

Subjects were able to keep a permanent visual record of their experimental outcomes because results of successive tribble plantings remained on the screen in the form of green or purple horizontal lines/dots for all eight trials (see Appendix C for sample data displays generated by subjects in Condition 4). After the eighth tribble had been planted, the computer screen presented the following written message: "WAIT, YOU HAVE 4 EXTRA TRIBBLES, 1 EXTRA CHECK." The display area was then extended downward to allow the subject space for four additional tribble plantings. This manipulation was included in order to generate the impression that experimental resources were limited, an impression that would incline subjects to engage in hypothesis testing rather than attempting to solve the problem using some kind of "pattern-matching" strategy. That is, it seemed plausible to suppose that subjects given an infinite number of tribbles might elect to plant all over the screen, find the line that best bisected green lines/dots from purple lines/dots, and locate the hypothetical moisture line at this point. Such a strategy requires little reasoning or forethought and would thus have undermined the purposes of the present investigation. The planting of the twelfth tribble marked the completion of the hypothesis-testing task for subjects. A complete copy of the computer program used to generate the task display is contained in Appendix B.

Dependent Variables Derived from Task

Three major dependent variables were assessed during the hypothesis-testing phase of the study:

1. The number of times in the course of 12 trials that the subject elected to move the hypothetical moisture line. Each movement of the moisture line was taken to represent a reformulation of the subject's current working hypothesis about the phenomenon under investigation. Failure to relocate the moisture line was thus interpreted as "inferential perseveration"--unwillingness to abandon a disconfirmed hypothesis.

2. The proportion of disconfirmatory outcomes which resulted in probe checks minus the proportion of confirmatory outcomes which resulted in probe checks, for each subject. A larger proportion of disconfirmatory trials resulting in probe checks would thus yield a positive value for this dependent measure, consistent with the hypothesis that subjects were more likely to question their methodology following outcomes which differed from expectation. (Recall that a disconfirmatory trial was one in which the subject made a prediction regarding whether the tribble would live or die which was not borne out by the data [the subject predicted that the tribble would live,

but it actually died, or vice-versa]. A confirmatory trial was one in which the subject's prediction regarding tribble life or death was consistent with the outcome on that trial.)

3. The distance between the subject's location of the line following trial 12 and the true location of the critical moisture line. This measure was intended to assess whether introduction of data error influenced the number of experiments required for identification of the actual critical moisture line.

Additional Dependent Variables

Following completion of the hypothesis-testing task, five additional dependent measures were obtained for subjects. The first two of these measures were intended to provide feedback regarding the impact of the experimental manipulations, to assess participants' subjective impressions of the ambiguity of the data which had been presented.

Confidence Measures. Subjects were asked to rate, on a 7-point scale, how confident they were that their final placement of the hypothetical moisture line was within one-half inch of the true location of the critical moisture line. Instructions for use of the 7-point scale, plus a copy of the scale itself, are contained on p. 117 of Appendix A.

Additional Tribbles. Subjects were asked how many additional tribbles they would need to plant in order to be absolutely certain that they had located the actual critical moisture line (see p. 117, Appendix A).

Memory. As soon as the task was complete, the experimenter erased the data record from the screen. Subjects in all four groups were then asked to estimate how many of their predictions were borne out by the data (p. 117, Appendix A). The difference between the actual number of accurate predictions and the estimated number of accurate predictions was computed for each subject. There was no prior empirical or theoretical foundation for inclusion of this dependent measure.

Subjective Data Assessment. Following collection of all quantifiable data, subjects answered a question regarding the "quality" of the data they had received (p. 117, Appendix A). This question was designed to provide additional, direct information regarding participants' subjective experiences of the different forms of data error.

Tape Recorded Comments. As will be recalled, during the hypothesis-testing phase of the study participants were asked to "think out loud" into a tape recorder about their thoughts, perceptions, and inferential strategies. These tape-recorded protocols were then transcribed and typed in such a manner that a given transcript could be matched with the computer record of that subject's input. It was hoped that these protocols would provide additional insight into the various inferential approaches elicited by inclusion of one or both types of error into the data returned to subjects.

CHAPTER V

Results

Overview

A complete presentation of the raw data generated by each subject is contained in Appendix C. What follow are statistical analyses and data summaries for each dependent variable in sequential order, beginning with quantitative results and concluding with qualitative data derived from participant self reports and visual inspection of the data.

Frequency of Hypothesis Revision as a Function of the Presence of Data Error

For data analysis, the number of experiments out of a total of twelve which resulted in hypothesis revision (relocation of hypothetical moisture line), was recorded for each subject. Descriptive statistics for this dependent variable are contained in Table 4. Results of a two factor (measurement x stochastic error) analysis of variance for this data are summarized in Table 5. The analysis revealed a statistically significant effect for the presence of stochastic error, $F(1,36) = 30.54$, $p < .001$. Subjects who received data using instrumentation which malfunctioned 25% of the time were significantly more hesitant to revise a disconfirmed hypothesis than were subjects whose feedback data did not incorporate this form of

Table 4

The Total Number of Hypothesis Revisions for
Each Subject in Conditions 1-4

Subject Number	No Error	Measurement Error Only	Stochastic Error Only	Both Measurement & Stochastic Error
1	10	9	1	2
2	10	5	1	3
3	8	10	4	3
4	5	9	3	7
5	8	5	3	8
6	9	7	3	3
7	8	7	5	0
8	5	7	5	5
9	4	7	5	5
10	8	4	2	5
Sum X	75	70	32	41
$\bar{X}$	7.5	7.0	3.2	4.1
S^2	4.49	3.76	2.40	5.66

Table 5

Analysis of Variance for Total Number of Hypothesis Revisions

Source	S.S.	d.f.	M.S.	F	p
Stochastic Error	124.60	1	124.60	30.54	.001*
Measurement Error	.40	1	.40	.10	n.s.
Stochastic x Measurement Error	9.90	1	9.90	2.43	n.s.
Error	147.36	36	4.08		

error. A magnitude of effect estimate (R^2) revealed that 44% of the variance in frequency of hypothesis revision was attributable to the introduction of stochastic error as an independent variable. The effect for measurement error was negligible ($R^2 = .0014$) and did not approach statistical significance, $f(1,36) = .10$, n.s. Moreover, the data did not yield a significant interaction effect when measurement and stochastic error were combined for subjects in Condition 4, $R^2 = .04$, $F(1,36) = 2.43$, n.s. Figure 8 shows the impact of stochastic error on subjects' willingness to abandon disconfirmed hypotheses. Mean number of moisture line relocations is plotted for each of the four groups, revealing a substantial decline in this variable resulting from the introduction of instrument malfunction.

It might seem plausible to suppose that the failure to obtain an effect for measurement error, particularly in the Measurement Error Only condition, can be ascribed to the unique parameters of the experimental task (and thus not generalizable to other hypothesis-testing contexts). Note that the generation of a dead tribble east of the hypothetical moisture line allowed the subject to conclusively reject his/her current location of the line in favor of a new location at the westernmost border of the planting range (see Figure 9). Such conclusive rejection eliminates any potential informativeness of measurement error, and therefore would not be expected to influence frequency of hypothesis revision. However, according to this reasoning, once a subject begins to test hypotheses within the vicinity of the true moisture line (where ambiguous experimental results are generated), frequency of hypothesis revision should decrease;

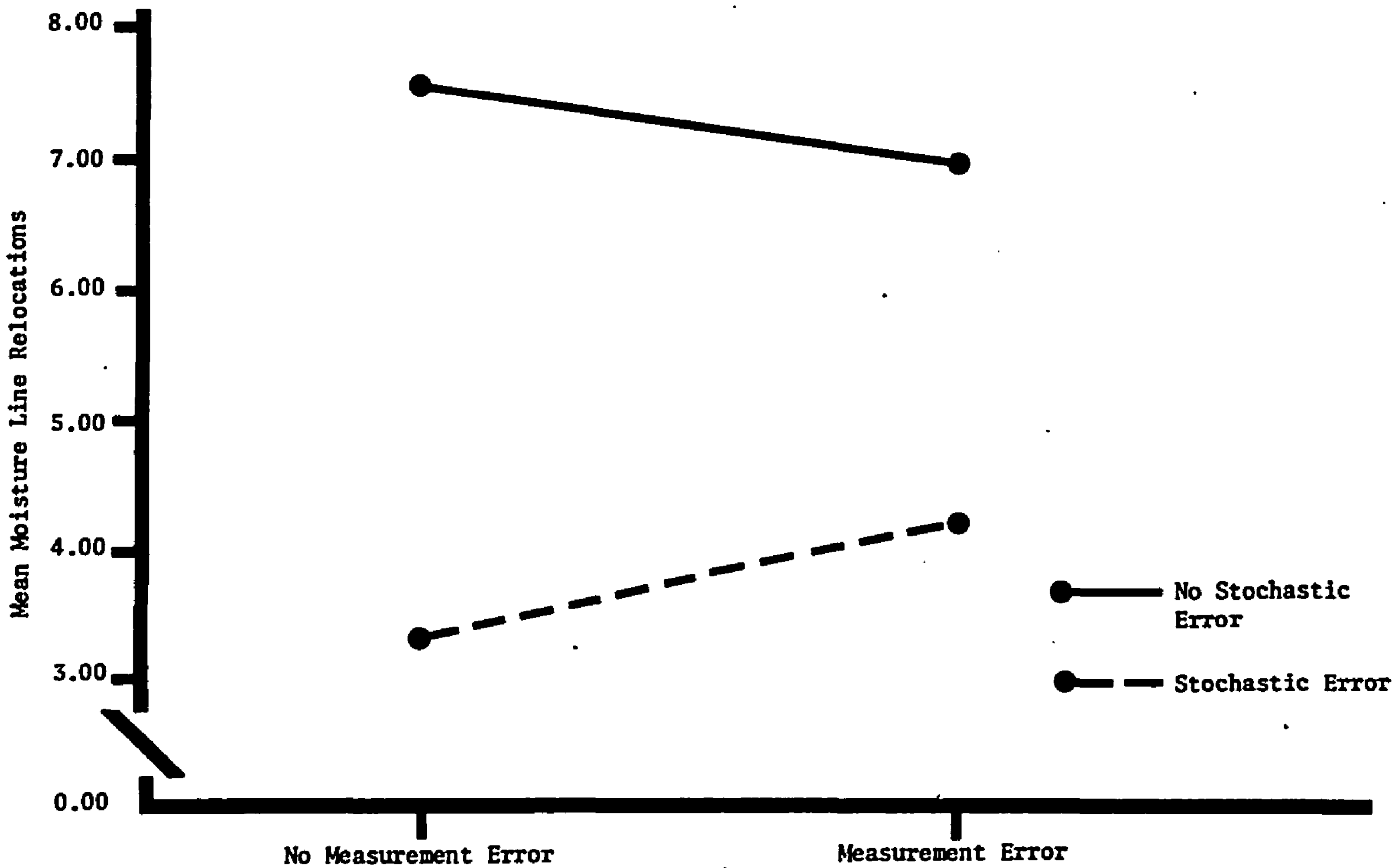


Figure 8: Mean number of moisture line relocations as a function of experimental condition.

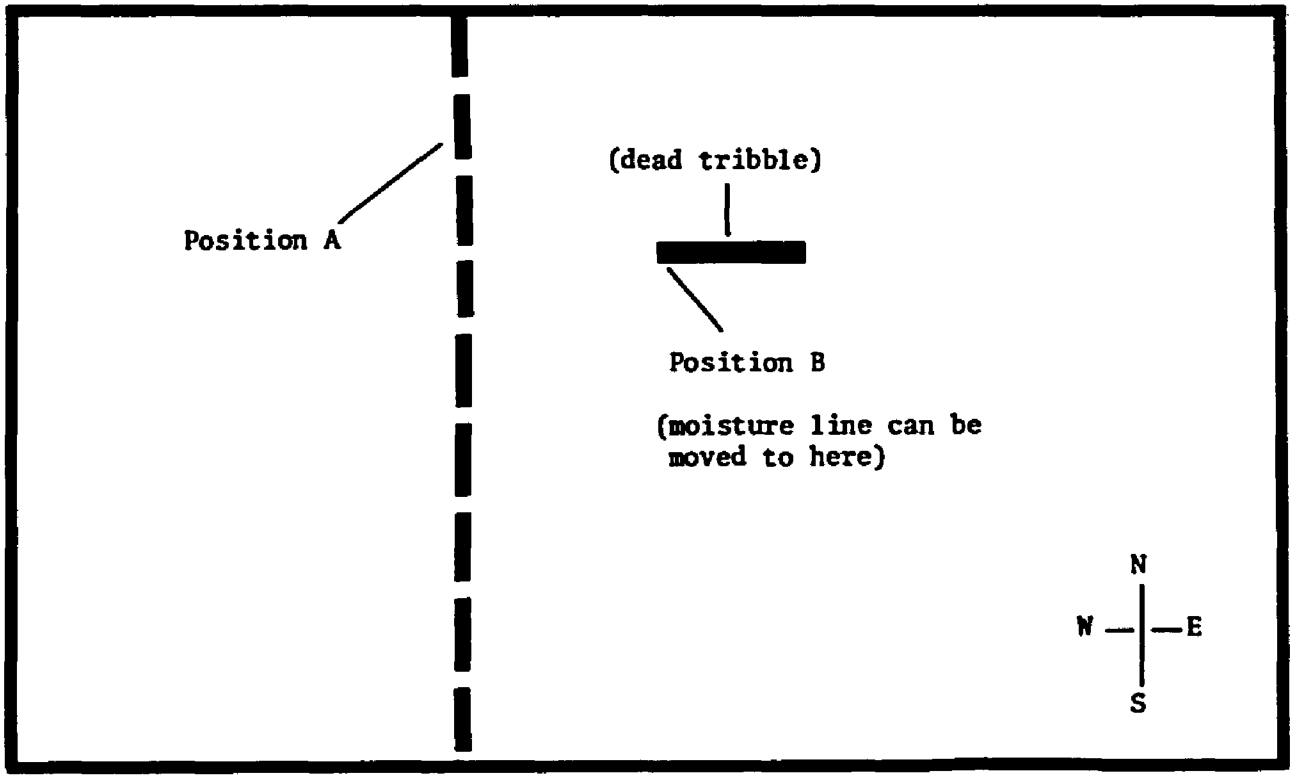


Figure 9: In the absence of stochastic error the hypothetical moisture line can be conclusively relocated from Position A to Position B.

unambiguous falsification is no longer possible. Because the majority of subjects in the Measurement Error Only Condition (6/10) reached the vicinity of the true moisture line within six trials, a comparison of the performance of subjects in Conditions 1 (no error) and 2 (measurement error only) for trials 7-12 should have yielded a difference with respect to frequency of hypothesis revision; subjects in Condition 2 should have revised less frequently than subjects in Condition 1 on these trials. In fact, subjects in Conditions 1 and 2 were equally likely to revise hypotheses on trials 7-12, $t(18) = 0.00$, n.s. Even in the face of conflicting or ambiguous data caused by the presence of measurement error, subjects are quite willing to replace one hypothesis with another which is no more conclusively supported by the data at hand.

Probe Check Likelihood Following Disconfirmatory, Versus Confirmatory, Trials

Because the frequency of disconfirmatory outcomes occurring for a given subject varied as a function of that subject's particular hypothesis-testing strategy, the absolute number of probe checks following disconfirmatory trials for each participant was not an appropriate index for use in the comparison of group means.[2] It was therefore necessary to compare the proportion of disconfirmatory outcomes which resulted in probe checks to the proportion of confirmatory outcomes which resulted in probe checks. A greater proportion

---

[2]Note that the No Error and Measurement Error Only Conditions are not relevant with respect to this dependent variable.

of probe checks following disconfirmatory than confirmatory outcomes would support the contention that experimental methodology comes under greater scrutiny following generation of unexpected data. Accordingly, the proportion of confirmatory trials resulting in probe checks was subtracted from the proportion of disconfirmatory trials resulting in probe checks for each subject (Table 6). Across subjects, the differences, although in the predicted direction overall, were only marginally significant, Wilcoxin's $T(20) = 65.5$, $p < .071$. Subjects were somewhat more likely to examine their methodology following outcomes inconsistent with their current working hypothesis about the location of the critical moisture line, than following consistent outcomes (Figure 10).

Confidence Ratings

Recall that subjects rated, on a 7-point scale, how confident they were that their final placement of the hypothetical moisture line was within one-half inch of the true location of the critical line. Descriptive statistics for responses to this question are presented in Table 7. A two-way analysis of variance of the data (summarized in Table 8) revealed that subjects exposed to stochastic error were significantly less certain that they had discovered the true location of the critical moisture line by the end of the task than were subjects whose data did not incorporate this form of error, $F(1,36) = 4.48$, $p < .05$. The magnitude of effect estimate for this finding ($R^2 = .10$) was substantially smaller than was the previously reported estimate for frequency of hypothesis revision. The effect of measurement error on confidence ratings was not statistically

Table 6

The Proportion of Disconfirmatory Outcomes Which Resulted in Probe Checks Minus the Proportion of Confirmatory Outcomes Which Resulted in Probe Checks for Subjects in Conditions 3 and 4

	Group 3		Group 4	
	Stochastic Error		Measurement and Stochastic Error	
Subject Number	Difference	Rank	Difference	Rank
1	-1/6	-8.5	+9/35	+12
2	-1/6	-8.5	+3/8	+17
3	+3/7	18	+1/6	+8.5
4	-1/6	-8.5	+3/35	+4
5	+3/6	+19	+3/9	+15
6	+2/7	+13	+3/5	+20
7	+3/35	+4	-1/3	-15
8	+0	+1.5	-2/6	-15
9	-0	-1.5	-1/6	-8.5
10	+1/6	+1.5	+3/35	+4

Sum of negative ranks = 65.5 = T

Sum of positive ranks = 144.5

Note. "Rank" refers to the rank of the difference values (based on the size of the number, beginning with the smallest) used in computing Wilcoxin's T statistic.

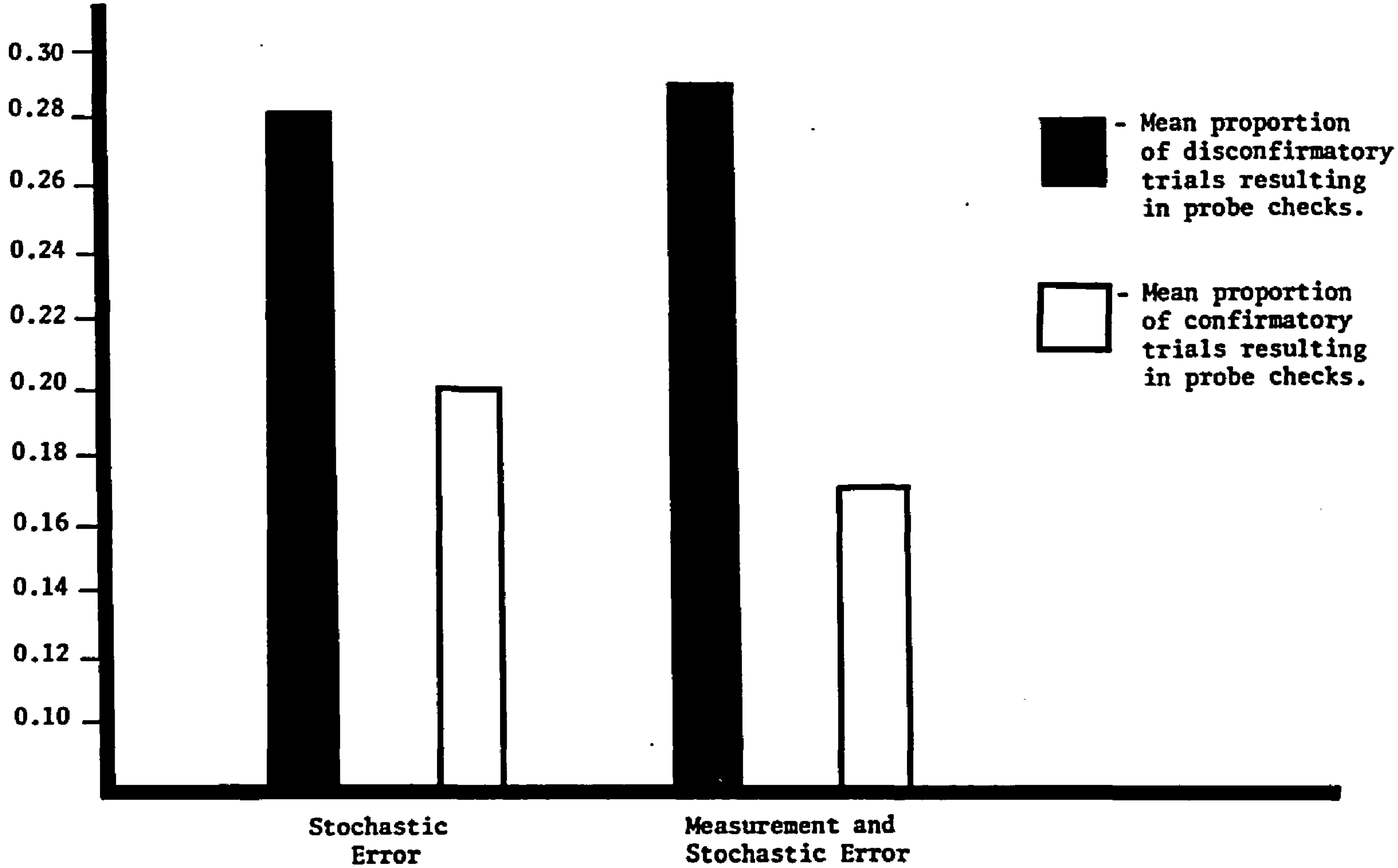


Figure 10: Mean proportion of disconfirmatory, versus confirmatory, trials which resulted in probe checks for subjects in Conditions 3 and 4.

Table 7

Confidence Ratings Regarding Final Placement of Hypothetical Moisture Line. Possible Range Equals 1 to 7, with 7 Indicating Extreme Certainty of Correct Placement

Subject Number	No Error	Measurement Error	Stochastic Error	Measurement & Stochastic Error
1	7	6	5	5
2	7	6	5	5
3	6	1	1	6
4	6	5	3	6
5	7	7	7	5
6	7	6	7	5
7	7	1	5	3
8	6	6	4	6
9	6	7	6	6
10	6	6	1	1
Sum X	65	51	44	48
$\bar{X}$	6.5	5.1	4.4	4.8
S^2	.28	4.97	4.71	2.62

Table 8

Analysis of Variance for Confidence Ratings

Source	S.S.	d.f.	M.S.	F	p
Stochastic Error	14.4	1	14.4	4.48	.05*
Measurement Error	2.5	1	2.5	.79	n.s.
Stochastic x Measurement Error	8.1	1	8.1	2.57	.20
Error	113.4	36	3.15		

*Note that the five percent level of significance might be slightly inflated, due to heterogeneity of variance.

significant, $F(1,36) = .79$, n.s., nor was the interaction for the two forms of error in combination, $F(1,36) = 2.57$, n.s. Figure 11 shows the mean confidence ratings for the four conditions.

Proximity to True Moisture Line at Completion of Task

To assess whether introduction of data error influenced the number of experiments required for identification of the actual critical moisture line, the distance between final placement of the hypothetical line and location of the true line was computed for each subject. These values (derived using the 280 equidistant units of space across the monitor of the Apple II) are listed in Table 9. The most striking aspect of this data is the heterogeneity of variance which emerged across conditions, F-max $(9,9) = 1210.39$, $p < .001$. The variance estimates for subjects exposed to the four conditions are as follows: (1) no error, $s^2 = .77$; (2) measurement error only, $s^2 = 43.16$; (3) stochastic error only, $s^2 = 824.84$; (4) both measurement and stochastic error, $s^2 = 932.08$. Although the introduction of measurement error produced a substantial increase in variance on this dependent measure, this increase was small in comparison to that which resulted from the introduction of stochastic error. It appears that stochastic error induces much more diversity in subjects' responses to incoming data, particularly with regard to likelihood of hypothesis revision following disconfirmation, than does measurement error.

Additional Tribbles Requested

Following completion of the task, participants were asked how many additional tribbles they would need to plant in order to be

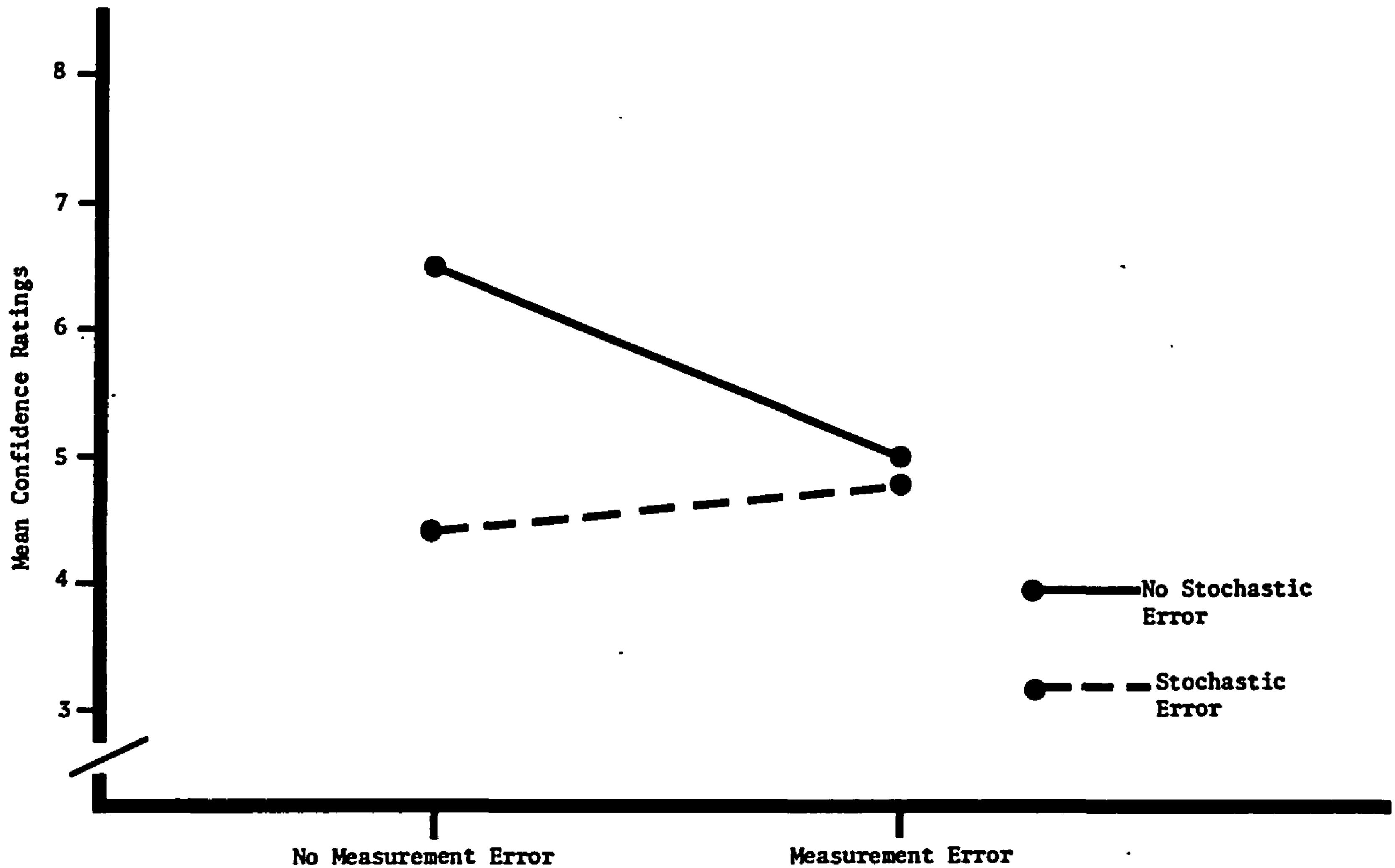


Figure 11: Mean confidence ratings for subjects in each of the four experimental groups.

Table 9

Descriptive Statistics for Proximity to True Moisture Line at Completion of Task

Subject Number	No Error: Final Location of Line		No Error: Absolute Value of Difference	Measurement Error Only: Final Location of Line		Measurement Error Only: Absolute Value of Difference
1	210	-210*	0	209	-210	1
2	211	-210	1	214	-210	4
3	207	-210	3	227	-210	17
4	209	-210	1	222	-210	12
5	211	-210	1	207	-210	3
6	211	-210	1	217	-210	7
7	210	-210	0	216	-210	6
8	211	-210	1	210	-210	0
9	209	-210	1	206	-210	4
10	212	-210	2	191	-210	19
			$\Sigma x = 11$			$\Sigma x = 73$
			$\bar{x} = 1.1$			$\bar{x} = 7.3$
			$s^2 = .77$			$s^2 = 43.16$
	Stochastic Error Only			Both Measurement & Stochastic Error		
1	202	-210	8	199	-210	11
2	131	-210	79	216	-210	6
3	209	-210	1	234	-210	24
4	206	-210	4	206	-210	4
5	195	-210	15	202	-210	8
6	208	-210	2	218	-210	8

Table 9 (Continued)

Subject Number	Stochastic Error Only: Final Location of Line		Stochastic Error Only: Absolute Value of Difference	Both Measurement & Stochastic Error: Final Location of Line		Both Measurement & Stochastic Error: Absolute Value of Difference
7	207	-210	3	110	-210	100
8	172	-210	38	210	-210	0
9	212	-210	2	218	-210	8
10	274	-210	64	255	-210	45
			$\Sigma x = 216$			$\Sigma x = 214$
			$\bar{x} = 21.6$			$\bar{x} = 21.4$
			$s^2 = 824.84$			$s^2 = 932.08$

*Represents value for true location of critical moisture line.

absolutely certain that they had located the actual critical moisture line. Responses to this question are displayed in Table 10. Note that, in all four conditions, there was a great deal of variability in subjects' answers to this question, F-max (9,9) = 34.092, $p < .001$. The greatest variance in number of additional tribbles requested occurred for subjects exposed to stochastic error only. Note however, that a substantial proportion of this variance is attributable to inclusion of a single score; one subject indicated that 500 tribbles would be necessary for definite determination of the critical level. When this score is removed from the data the variance estimate reduces from 26,461.53 to 8,495, still considerably higher than for the other three groups, F-max (9,9) = 10.95, $p < .01$.

Heterogeneity of variance precluded the possibility of a test for significant differences between means for the various groups. As might be expected from the large number of additional tribbles requested by the subject in the stochastic error only condition, it was that condition which showed the highest mean request. Median number of additional tribbles requested by subjects in Conditions 1 through 4 were as follows: (1) no error, 15; (2) measurement error only, 20; (3) stochastic error only, 8; (4) measurement and stochastic error, 22.5.

Memory Test

Subjects were asked to estimate, from memory, how many of their 12 experiments had yielded predicted outcomes. The difference between the estimated number of accurate predictions and the actual number of accurate predictions was computed for each subject. These

Table 10

Number of Additional Tribbles Requested Following Completion of Task

Subject Number	No Error	Measurement Error	Stochastic Error	Measurement & Stochastic Error
1	100	-	8	20
2	15	0	500	15
3	-*	20	30	100
4	0	20	30	25
5	12	100	0	30
6	2	0	4	25
7	0	2	5	8
8	15	50	20	25
9	20	20	-	3
10	18	20	8	4
Σx	182	232	605	255
$\bar{x}$	22.75	29	67.22	25.5
s^2	1025.92	1053.65	26,461.53	776.18

*Indicates that subject failed to answer this question or gave an answer that was uninterpretable, such as, "Would need many, many more tribbles."

difference scores are presented in Table 11. Subjects in the no error and measurement error only conditions showed a slight mean tendency to underestimate their predictive accuracy, while subjects in the other two groups (stochastic error only and stochastic and measurement error combined) tended to overestimate their predictive accuracy by approximately 1.5 outcomes (Figure 12). Results of a two-way analysis of variance of the data (summarized in Table 12) indicated a statistically significant effect for stochastic error, $F(1,36) = 10.45$, $p < .005$, $R^2 = .22$. The effect for measurement error did not approach statistical significance, $F(1,36) = .0025$, n.s., nor did the interaction effect, $F(1,36) = .53$, n.s. Thus, the introduction of stochastic error appears to have inflated "recall" of prediction accuracy.

Discipline Differences

To assess whether participants in each of the three scientific disciplines may have responded differently on any of the dependent measures, means and variances for each dependent variable were computed according to discipline, ignoring all independent variables. These values are presented in Table 13. The only means which differed substantially according to discipline were those for probe check likelihood following disconfirmatory, versus confirmatory, trials. Chemistry students were more likely to conduct probe checks following disconfirmatory outcomes than were participants from the other two disciplines. Otherwise, there appeared to be no systematic differences in the dependent measures attributable to scientific/ discipline.

Table 11

The Difference Between the Estimated Number of Accurate Predictions and the Actual Number of Accurate Predictions for Subjects in Conditions 1 Through 4

	No Error			Measurement Error		
Subject Number	Estimated Predictions Correct	Actual Predictions Correct	Difference	Estimated Predictions Correct	Actual Predictions Correct	Difference
1	8	8	0	6	6	0
2	6	6	0	6	8	-2
3	4	4	0	3	4	-1
4	7.2	9	-1.8	5	7	-2
5	6	6	0	6	6	0
6	3	4	-1	7	6	1
7	3.6	10	-6.4	4	4	0
8	6	6	0	5	5	0
9	9	8	1	6	7	-1
10	4.8	6	-1.2	8	6	2
			$\Sigma x = -9.4$			$\Sigma x = -3$
			$\bar{x} = -.94$			$\bar{x} = -.3$
			$s^2 = 4.33$			$s^2 = 1.56$
	Stochastic Error			Measurement and Stochastic Error		
1	9	6	3	6	7	-1
2	6	6	0	8	4	4
3	7.2	5	2.2	6	6	0
4	6	6	0	8	5	3

Table 11 (Continued)

Subject Number	Stochastic Error: Estimated Predictions Correct	Stochastic Error: Actual Predictions Correct	Stochastic Error: Difference	Measurement and Stochastic Error: Estimated Predictions Correct	Measurement and Stochastic Error: Actual Predictions Correct	Measurement and Stochastic Error: Difference
5	10	6	4	3	3	0
6	6	5	1	9	7	2
7	6	5	1	9	9	0
8	10	4	6	10	6	4
9	8	8	0	7	6	1
10	6	6	0	6	5	1
			$\Sigma x = 17.2$			$\Sigma x = 14$
			$\bar{x} = 1.72$			$\bar{x} = 1.4$
			$s^2 = 4.24$			$s^2 = 3.17$

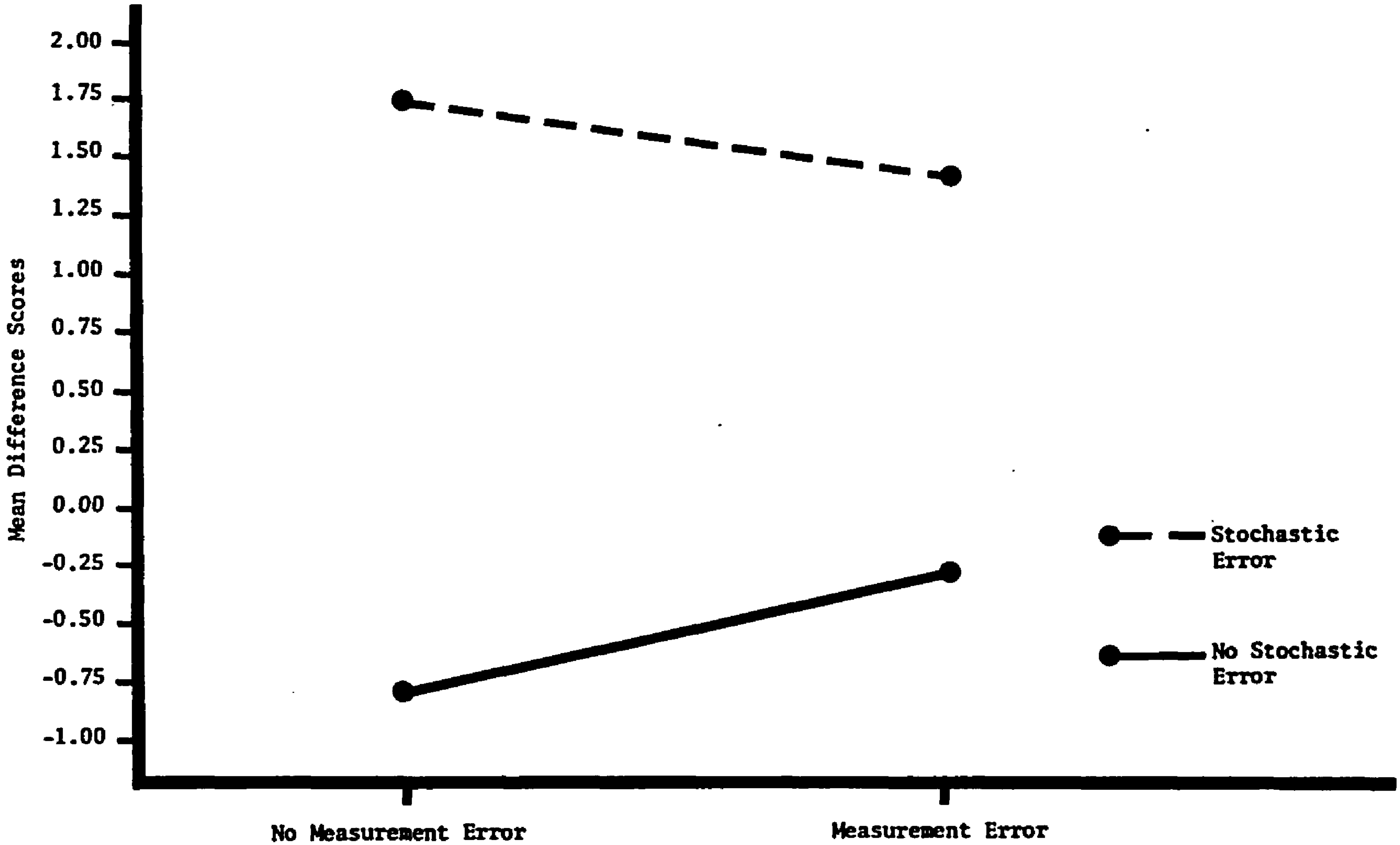


Figure 12: Mean difference scores between estimated and actual number of accurate predictions. A difference score of 0 indicates perfect recall; negative scores indicate a tendency to underestimate prediction accuracy; positive scores indicate a tendency to overestimate prediction accuracy.

Table 12

Analysis of Variance on Difference Between Estimated and Actual Number of Accurate Predictions

Source	S.S.	d.f.	M.S.	F	p
Total					
Stochastic Error	47.52	1	47.52	14.31	.001
Measurement Error	.25	1	.25	0.08	n.s.
Stochastic x Measurement Error	2.31	1	2.31	0.70	n.s.
Error	119.56	36	3.32		

Table 13

Means and Variances for Each Dependent Measure,
Computed According to Discipline,
Ignoring All Independent Variables

Frequency of Hypothesis Revision

	Physics	Chemistry	Biology
$\bar{X}$	5.63	5.50	5.17
s^2	11.16	7.73	2.53

Probe Check Likelihood Following Disconfirmatory, Versus Confirmatory Trials

	Physics	Chemistry	Biology
$\bar{X}$	.1016	.2452	-.0413
s^2	.0609	.1121	.0328

Confidence Ratings

	Physics	Chemistry	Biology
$\bar{X}$	5.00	5.58	5.08
s^2	3.35	3.72	4.08

Proximity to True Moisture Line at Completion of Task

	Physics	Chemistry	Biology
$\bar{X}$	11	12.83	15.33
s^2	372.88	771.17	472.19

Additional Tribbles Requested

	Physics	Chemistry	Biology
$\bar{X}$	58.62* 21.83	16.09	25.55
s^2	18,241.20* 711.82	839.84	160.28

Table 13 (Continued)

	Memory Test		
	Physics	Chemistry	Biology
$\bar{X}$	.28	0.13	1.07
s^2	3.39	5.90	4.33

*Note that this mean and variance were computed twice, the first value includes the score of 500 for subject #12, the second was computed disregarding this score.

Qualitative Results

Involvement With Task. In general, subjects appeared to become rapidly engrossed in the task and seemed highly motivated to locate the actual critical moisture line. Considerable time was spent in formulating predictions which subjects believed would generate maximum information yield. A few subjects spent seven to ten minutes in selecting each planting location such that 1 1/2 - 2 hours were required for task completion. Not a single subject appeared to carelessly conduct 12 experiments in rapid succession, meeting only the minimal requirements for earning the $10.00 payment. In fact, three subjects insisted that they should not be paid because they had found the task so enjoyable (they were paid, nevertheless).

Subjects' Assessment of the Feedback Data. Following completion of the task subjects were asked to describe, in writing, how "good" they felt the data they had received was. They were asked to describe how "diagnostic/helpful/valid/differentiating" it was. Themes which emerged in response to this question are exemplified by the following statements from subjects in each of the four conditions.

1. NO ERROR CONDITION. Subjects in this group generally indicated that they felt the data quality was good, that it permitted conclusive inferences about the location of the critical moisture line.

> --The data was very conclusive. If the plant lived or died it was obvious if the line belonged either east or west of that location. In this respect the data quality was excellent.

--The responses (data) indicated a good range for the life or death of the tribble. Within close approach of the line, the responses differentiated where the line was.

--The data showed quite consistently where the growth line existed since several trials were made about the position. Enough data was taken to make the experimenter certain that the results were as accurate as possible.

2. MEASUREMENT ERROR ONLY. Only one theme emerged in the responses of subjects exposed to Condition 2. A number of participants indicated that measurement error had impeded their ability to draw conclusive inferences (despite the fact that they revised their hypotheses frequently) when they began to hypothesis test within close proximity of the true moisture line.

--The data was very good for the first few trials when I was trying to find the location of the life line on a large scale. As the range of possible locations narrowed, however, each individual planting gave very little additional information.

--The data was helpful in the coarse adjustment of the line, but when it came to the fine adjustment of the line, the lack of a Gaussian distribution confused me slightly and affected my strategies. A Gaussian distribution would have been more differentiating.

--It was very good with respect to narrowing the possible range down to within plus or minus one-half inch. Beyond that, it became extremely difficult to find the exact location of the line, due to the uniform probability of the tribble being anywhere within the range represented by the horizontal line.

3. STOCHASTIC ERROR ONLY. Subjects exposed to stochastic error indicated that they found the data frustrating to work with because it was so difficult to "trust" any given data point.

> --There was a problem with assuming the data to be "good" because of the possibility of an unreliable point. I felt the data provided an easy way to set a maximum or minimum line, but it was hard to "trust" it enough to feel that you had very accurately located the line.

> --The data was helpful with the qualification that 1 out of 4 results may be invalid, and the additional hindrance to determining the correct answer, of a limited number of probe correctness determinations. This is frustrating and unlike the real situation in biology where results should be able to be consistently repeated.

> --The data I received was not as diagnostic as I'd have preferred, particularly because the probe checks were only useful when "valid"; when unreliable, it was still guesswork as to the true status of my tribble. I could check this, however, by planting two tribbles at the same moisture level and then comparing their survivability. If, for example, they both lived, then that moisture level was probably tribble supportive.

4. BOTH MEASUREMENT AND STOCHASTIC ERROR. Although subjects in this condition indicated that they were aware of the presence of both measurement and stochastic error in their data, the primary focus of their responses was on the difficulty of interpreting empirical results due to the possibility of "probe malfunction."

> --I didn't appreciate the lack of reliability of the "equipment." It is difficult to differentiate between data points when there is only a 75% chance of each point being reliable. I'd rather redo the experiment with better equipment.

--I think I would have put some checks on the information I was reading from the probe, an efficient way of making sure data was o.k. I thought the data was lousy and would want to be able to do better, it was helpful, however, in that it got me somewhere.

--With an average reliability of only three valid data points for every four, I would say the data quality was not very good. My data seemed to represent this statistical criterion, based upon the hypothetical line. The probe checks would be absolutely necessary for the determination of the actual location of the line.

--The quality was not as good as I would have liked. (1) There was an equal chance that the probe would land anywhere within a 15 mile radius and (2) the data received was correct in only approximately 75% of the trials with a limited number of probe checks.

Response Trends. Using the raw data contained in Appendix C, supplemented by subjects' recorded verbal protocols, it was possible to tentatively identify trends in the hypothesis-testing strategies employed by subjects across the four experimental conditions. A brief discussion of these trends follows.

1. NO ERROR CONDITION. Subjects appeared to be quite consistent in the strategies they employed when the data was error free. In general, two inferential approaches were predominant in these data. The first involved planting a tribble immediately to the right of the hypothetical moisture line; when this tribble died, the line was moved from its original location to a point just east of the dead plant. Trial 2 was a repetition of this strategy--plant to the east of the line, when the tribble dies, relocate the moisture line

eastward in accordance with the new data. This sequence was repeated until a tribble planted to the east of the hypothetical moisture line lived. If any space remained between this location and the location of the last dead tribble, the strategy was employed in reverse (using smaller increments) until an exact location for the true moisture line could be specified (see Appendix C, Subject 3, Condition 1 for the raw data from one participant which epitomizes this strategy).

The second strategy which tended to be employed by subjects in this condition involved combining the method already described with another method, analogous to that used in finding a solution to the children's game "hi-lo" ("I'm thinking of a number between 1 and 100..."). That is, the subject would locate two tribbles, separated by some distance on the east-west dimension, such that one was living and the other was dead. On the third trial this distance would be bisected by locating a tribble exactly halfway between the first two. Whether this third tribble lived or died determined the next distance to be bisected such that each trial yielded successively closer approximations to the location of the true moisture line. As already indicated, this "successive approximation" strategy tended to be combined with the earlier "incremental" method of hypothesis testing such that the typical subject tested incrementally for the first few trials and then replaced this strategy with the more efficient "hi-lo" method later in the task (see Appendix C, Subject 1, Condition 1 for an example). Although it's impossible to be certain, it appears that the incremental method was used exclusively by 2 subjects (numbers 2 and 3), while the combination method was employed by the remaining 8.

2. MEASUREMENT ERROR ONLY. The strategies employed by subjects in Condition 2 appeared to be virtually identical to those used by the no-error subjects, except that once the subject began testing within the vicinity of the true moisture line it became more difficult to infer, from visual inspection of the data, what hypothesis-testing method was being employed. Hypothesis revisions within this region seemed to reflect the subject's "best guess" about appropriate moisture line relocations, since the conclusive inferences typical of subjects in the No Error Condition were impossible. At times it was difficult to infer a rational motive to the relocations which some subjects undertook in this area (see, for example, Appendix C, Subject 7, Condition 2, relocation following Trial 11). In general, it is interesting to note how apparently similar the inferential strategies used by subjects in Condition 2 are to those employed by subjects whose data did not contain measurement error. The lack of an effect for measurement error seems to emerge in the qualitative, as well as the quantitative, data.

3. STOCHASTIC ERROR ONLY. Unlike subjects in the Measurement Error Only Condition, subjects in the Stochastic Error Only Condition did not rely on a single predominant hypothesis-testing strategy. Some appeared to be extremely "conservative" in that they generated numerous pieces of redundant disconfirmatory data before revising the location of the hypothetical moisture line (see raw data for Subjects 2, 4, and 10, Condition 3). Others appeared to use a more "risky" strategy, making large revisions on the basis of a single piece of data (see data for Subject 6, Condition 3). In general, it

seems plausible to suppose that the two strategies used by subjects in the first two conditions were also employed by subjects in Condition 3, but were overlaid by large individual differences in how much redundant data needed to be generated prior to rejection of a current hypothesis.

There was some tendency toward regularity in the planting strategies subjects used after finding that one of their probes had malfunctioned. All together, 10 of the 29 probe checks undertaken by subjects in the Stochastic Error Only Condition (one subject decided to do only 2 probe checks) indicated a broken probe. Seven of these ten trials were followed either by replanting in the same location (n = 2) or by "backing up" to plant in a direction opposite that in which the subject had been moving (n = 5; usually switching from eastward to westward).

4. MEASUREMENT AND STOCHASTIC ERROR. As with subjects in the Stochastic Error Only Condition, it was difficult to identify regularities in the hypothesis-testing strategies of subjects exposed to both measurement and stochastic error. While some subjects seemed to rely on a combination of the "incremental" and "successive approximation" strategies described above, there were sufficient idiosyncracies in planting behavior across subjects to preclude the identification of definite regularities in the hypothesis testing behavior of these subjects. With respect to the discovery of a broken probe (total n = 8), however, subjects in the Fourth Condition, as with subjects in Condition 3, tended to follow this discovery with a repetition of the same experiment (n = 3) or with an

experiment which involved planting in a direction opposite that in which the subject had been moving previously (n = 2).

Recorded Protocols. Unfortunately, the written protocols derived from participants' verbalizations during the task did not permit a great deal of additional insight into the inference strategies subjects used across the four conditions. In a large part this was due to the fact that subjects became quite engrossed in the task and then tended to mumble disconnected phrases to themselves or to forget altogether (despite frequent prompts) that they should be talking out loud. Transcribed protocols from such subjects were thus choppy and difficult to follow for extended periods.

In general, the handwritten protocols tended to corroborate the impressions regarding hypothesis-testing strategies which were derived from visual inspection of the data. Numerous subjects made reference to planting "exactly between the last two" or to "moving by increments" across the planting area.

There was also some evidence that subjects believed it was somehow more "logical" or informative to do a probe check following a disconfirmatory trial, than following a confirmatory trial. A few direct quotations serve to illustrate this.

1. Predict results... Die. And it lived. At this point I think I would need to know if that was good data, so I'm going to use up one of my two probe checks.

2. Okay, that result confirms my ideas so it's obviously correct and it's a waste to check it. No, I do not want to check it.

3. No. I don't want to do a probe check now; not until I get a disconfirming experiment. Then it will be useful to check the probe.

CHAPTER VI

DISCUSSION

Overview

Consistent with presentation of data analyses thus far, the implications of significant results will be discussed in sequential order, following the serial presentation of dependent variables introduced in Chapter V. Discussion of the data will be followed by a consideration of limitations of the present investigation, and the Chapter will conclude with a brief speculative section regarding possible directions for future research.

Frequency of Hypothesis Revision

Stochastic Error. Perhaps the most striking result with respect to the effect of data error on frequency of hypothesis revision was the impact of stochastic error on this dependent measure. Subjects whose feedback data incorporated the concept of instrument malfunction were extremely hesitant to alter a disconfirmed hypothesis before considerable disconfirmatory data had been accumulated. Subjects appeared to behave as though none of the data they received were trustworthy, as though any given data point had a high probability of being unreliable, when in fact an average of 75% of the data was "known" to be legitimate.

Although the implications of this result will necessarily vary depending on the specific inferential task to which they are applied, one general interpretation of the data is as follows: Whenever an experimental hypothesis test yields disconfirmatory results which could plausibly be attributed to instrumentation failure, numerous replications may be undertaken before rejection and/or reformulation of the hypothesis is considered. Although there is no existing empirical literature which corroborates this interpretation, a preliminary application of the laboratory finding to a historical case study may serve to suggest the generalizability of the present result.

In 1831 Michael Faraday undertook a crucial experiment which led to the discovery of electromagnetic induction (the creation of an electromagnetic field in a conductor by another, nearby, electromagnetic field). Once this initial successful experiment was completed, Faraday began a series of further experiments intended to extend the theoretical implications of the initial finding. These experiments, which he carefully recorded in diary form (Faraday, 1932-36), have recently been systematically explored by Tweney (1981) in order to elucidate the cognitive processes underlying Faraday's research during the hundreds of experiments which flowed from the initial crucial finding. Specifically, Tweney focused on the role of confirmation and disconfirmation in Faraday's work in an effort to determine whether there were systematic differences in his response to these two forms of evidence. In all, 134 experiments carried out by Faraday between August 29, 1831 and November 3, 1831 were analyzed. While this analysis revealed considerable evidence of confirmation bias at a number of junctures in Faraday's inference process, one manifestation of this bias is of particular interest with respect to the results of the present research. Tweney notes that Faraday made repeated attempts to produce an induced current using ordinary magnets (as opposed to electromagnets). Despite six failures to detect the desired effect, Faraday persisted (making slight instrumentational adjustments) until he obtained a single confirmatory result, at which point he abandoned this problem in order to undertake 20 experiments directed at other hypotheses. In general, very few of Faraday's disconfirmed hypotheses were abandoned or modified. Rather, they were typically retested using different methodology until a confirmatory result was obtained. To quote Tweney directly:

> Some hypotheses are so central to his theoretical framework that he insists upon repeated confirmatory attempts--again and again he tried to find the result he wanted, rather than the result he got. In most of these he succeeded, eventually, in getting confirmation. Then, usually, he quit after the first confirmation and turned to something else. (p. 3)

The theoretical significance of the finding that stochastic error inhibits hypothesis revision may be increased by permitting speculation which goes beyond consideration of auxiliary assumptions related exclusively to instrumentation failure. Recall that in Chapter III it was suggested that an analogy might be drawn between stochastic data error and a false theoretical auxiliary assumption in that, when either is conjoined with a true hypothesis, predicted experimental outcomes will fail to occur despite the veracity of the hypothesis. What an individual scientist does (inferentially) upon discovering that an auxiliary assumption is false will vary greatly depending on whether the auxiliary relates to faulty instrumentation or to erroneous theoretical presuppositions, but the initial outcome is the same: a predicted empirical result has failed to occur. The investigator is faced with a choice: "Which shall I question, the truth status of my auxiliary or the truth status of my hypothesis?" It may be less cognitively demanding to attribute a failed experimental outcome to a false auxiliary assumption than to undertake the mental work involved in changing one's conception of the world as is required in the reformulation of an experimental hypothesis. Moreover, given that it may be quite rational and appropriate to question the truth status of auxiliaries under many circumstances (witness the success of Faraday's research), it becomes even more plausible to suppose that the typical strategy of theory testing may involve the following sequence of events: (1) predict outcome; (2) observe results; (3) if results are disconfirmatory, examine auxiliaries; (4) if auxiliaries are all determined to be true, revise hypothesis; (5) if a false auxiliary is found, make appropriate instrumentation adjustment or peripheral conceptual revision such that essential aspects of theory are maintained. At an abstract level this appears to be a straightforward, efficient and methodical approach to theory testing, but it is interesting to consider whether there might be parameters of the inference task which enhance or inhibit the efficacy and appropriateness of this strategy. When an experimenter has developed a well articulated and highly corroborated theory such that auxiliary assumptions conjoined with any given focal hypothesis are few in number and entirely explicit, then the strategy outlined above might indeed provide an efficient means for assessing the truth status of a given theoretical formulation. However, consider the case in which an investigator is examining an ill-defined theory, all hypotheses derived from which are necessarily conjoined with innumerable auxiliary assumptions, each as ill defined as the theory itself (recall the example regarding a hypothesized association between compulsivity and susceptibility to post-hypnotic suggestion presented in Chapter II). Assume also that the investigator is not indifferent with respect to the outcome of a given empirical test but rather that the theory under investigation is one in which the researcher has considerable investment (cf. Mitroff, 1974). In an instance such as this, a basic cognitive bias toward avoidance of theoretical reformulation through prior examination of auxiliaries might interact with such a motivational tendency to permit the prolonged pursuit of a theoretically fruitless or patently false hypothesis. Thus, one highly speculative implication of the finding that stochastic error reduces the frequency of hypothesis revision is that auxiliary assumptions, both theoretical and "instrumentational," may promote a tendency to perseverate with the empirical test of a hypothesis even in the face of disconfirmatory data, sometimes resulting in the appropriate modification of peripheral aspects of a basically sound theory, and sometimes resulting in inappropriate perpetuation of a false theoretical conceptualization.

Implications for Clinical Psychology. With respect to the implications of the present data for clinical inference in psychotherapy, one is reminded of the work by Rosenhan (1973) suggesting that once a diagnostic hypothesis is formulated by the clinician (e.g., schizophrenia) a great deal of evidence inconsistent with this hypothesis (e.g., appropriate verbalizations, the absence of further hallucinations and delusions, excellent premorbid adjustment) may fail to result in the clinician's abandonment of the diagnosis due to what might be interpreted as a "retreat into auxiliaries" (e.g., the individual suffers from schizophrenia in remission). It is interesting to speculate whether clinicians entertaining erroneous diagnostic hypotheses are particularly immune to evidence which would tend to refute such hypotheses, due to the fact that numerous auxiliary assumptions are readily accessible which can effectively "explain away" disconfirmatory data. A few examples serve to illustrate: (1) The patient has attained sophistication with regard to staff expectations such that, while visual hallucinations are still experienced, they are no longer reported by the patient; or (2) Although the reinforcement history doesn't suggest that maternal attention was more frequent following temper tentrums on the part of the child, behavioral self reports are often inaccurate and thus should be discounted in this case; or (3) Although the results of neuropsychological assessment suggest that intellectual abilities have not deteriorated within the past year, the patient has been tested on so many previous occasions that scores may be inflated by a practice effect. The list is endless. The essential point is that, due to the ready availability of alternative explanations for disconfirmatory findings in clinical diagnosis, psychotherapists may have a special responsibility to entertain alternative hypotheses with respect to a given body of empirical data such that premature "freezing" of hypotheses (Wiggins, 1973) is minimized. Similarly, clinicians might guard against pursuit of false diagnostic hypotheses by deriving predictions from such hypotheses which would be relatively unlikely to occur unless the hypothesis under test were true. For example, the therapist who believes a child's temper tantrums are the result of reinforcement in the form of attention from the mother, might formulate the following empirical prediction: Effective control of maternal attention such that it never follows tantrum behavior should produce a measurable decrease in such behavior. Because this "experimental" outcome would be unlikely to occur unless the hypothesis under test were true, the clinician has made some attempt to avoid the trap of explaining away inconsistent data by an endless retreat into auxiliaries.

Measurement Error. One interesting aspect of the data with respect to the manipulation of measurement error was the absence of a detectable effect for this variable. When the data became conflicting or ambiguous due to the presence of measurement error, subjects remained quite willing to revise the experimental hypothesis with a frequency equal to that for subjects whose data did not incorporate this form of error. Visual inspection of the data for such subjects suggests that a number of these revisions may have been random in nature or were, perhaps, based on idiosyncratic "hunches" about the location of the true moisture line which were in no way conclusively implied by the available data (verbal protocols, as indicated in Chapter V, were of little help in this regard). It must be noted at this point, however, that while many of the relocations indicated by subjects under these circumstances appeared to be random, others may have been quite systematic and attributable to an artifact in the experimental methodology. Note that when a planting range containing a live tribble overlaps somewhat with a planting range containing a dead tribble (see Experiments 1 and 2, Figure 13), the subject in the Measurement Error Only Condition can deduce conclusively that the true location of the moisture line must lie between the leftmost end of the range containing the dead tribble and the rightmost end of the range containing the live tribble (refer again to Figure 13). Frequently, the subject confronted with this situation elected to place the hypothetical moisture line midway between the endpoints of the two overlapping ranges (although statistically it was equiprobable that the critical moisture line was at any one of the other possible locations between these two endpoints due to the uniformity of the landing distributions). When a third tribble was planted within the critical range already defined by the first two experiments, it was possible to narrow this potential range somewhat such that, as in the case depicted in Figure 13, the subject could be certain that the true critical line lay somewhere between the endpoints of the two most proximal ranges containing a live and a dead plant, respectively (in this case resulting from Experiments 1 and 3). This further narrowing of the critical range was often accompanied by a relocation of the hypothetical moisture line such that it was moved from its former location at the midpoint of the wide range defined by Experiments 1 and 2 to the midpoint of the narrower range

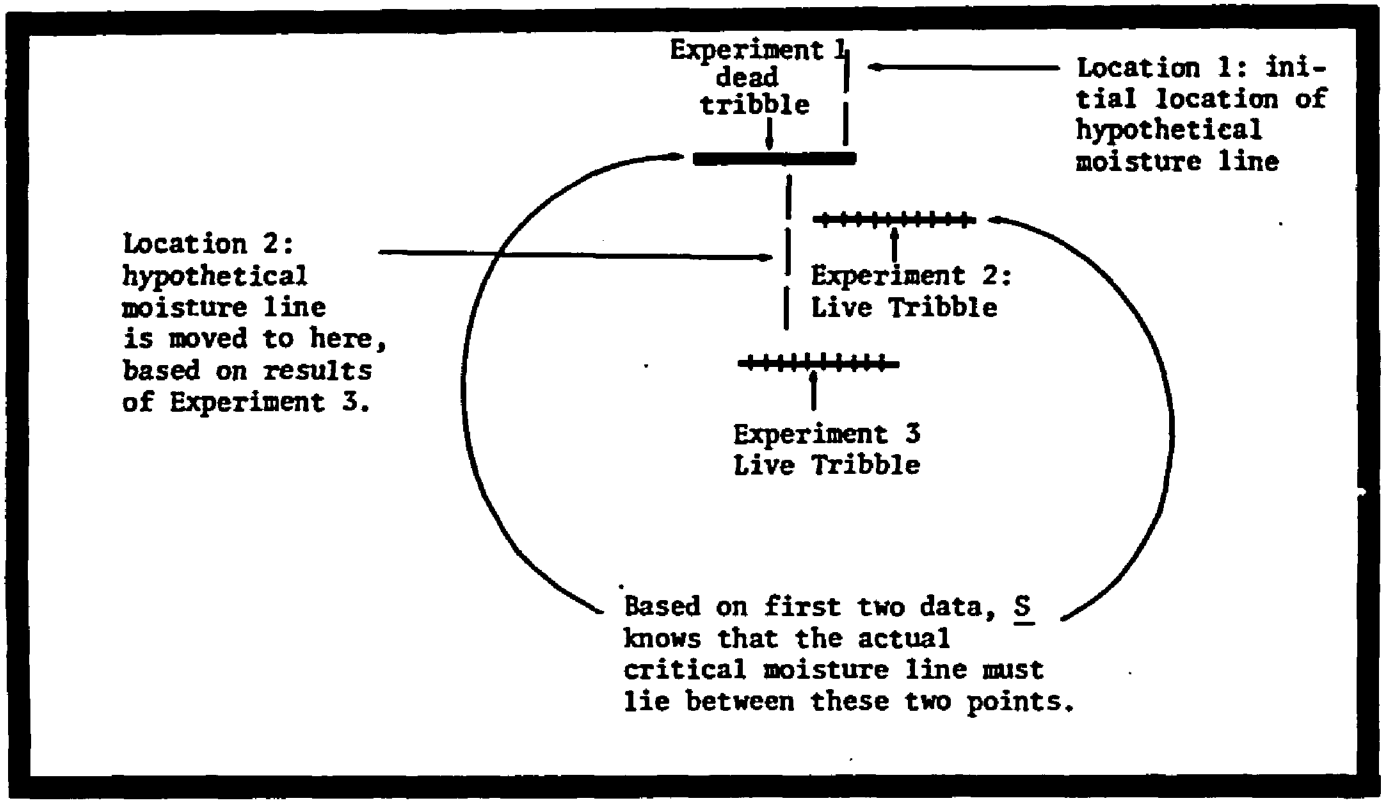


Figure 13: Possible explanation for continued hypothesis revision on Trials 7-12 by subjects in Measurement Error Only Condition.

defined by the inclusion of Experiment 3. Although there is no statistical basis for such a relocation, this type of response was not at all uncommon in the behavior of a number of subjects. One plausible explanation for such data relates to a point made in Chapter II: It is difficult for subjects to suspend beliefs which they hold about the natural universe simply because an experimenter has instructed them to do so. Recall that subjects exposed to measurement error were told that tribbles, although they could be aimed at a precise location, would not necessarily land where they were aimed but rather would be distributed uniformly throughout a specified landing range. Genuine cognitive integration of this description of the experimental task required that the subject suspend belief in the laws of gravitation which imply that plant landings would tend to be normally, rather than uniformly, distributed under such circumstances. Although subjects may have been superficially aware that the distributions within each planting range were uniform, they may have unwittingly been influenced by basic beliefs associated with notions of central tendency in the relocations they undertook on Trials 7-12.

Assuming for the purpose of speculation that measurement error does not affect frequency of hypothesis revision, what implications might this have for scientific hypothesis testing in general? Are there instances in which individuals' reasoning should be affected by the presence of measurement error, but it is not? While subjects in the present investigation know with certainty that the variance in each datum was attributable totally to measurement error, there are numerous instances in the actual practice of science where that is not the case. That is, variance which is attributed initially to measurement error may, at times, contain systematic variance attributable to some uncontrolled or unidentified variable unknown to the experimenter. A cognitive tendency to attribute disconfirmatory findings to measurement error may make the detection of such systematic variance less likely. When the detection of such variance would have significant theoretical implications for a particular branch of science, a disregard of "measurement" error may substantially slow the progress of discovery.

By drawing on a recent conversation with an experimental physicist regarding his own research, I can provide an example of a significant research question provoked by the detection of systematic patterns in what had previously been attributed to random fluctuations in data resulting from imprecision in measurement instrumentation. Newton's law of gravitation asserts that the gravitational force between two objects is proportional to the inverse of the square of the distance between the two objects. Measurements of the magnitude of gravitational force between objects of various sizes (ranging from atomic particles to planetary bodies) and at various distances of separation (ranging from angstrom units to light years) have yielded striking empirical conformity with the inverse square law. But periodic deviations from the values predicted by the law have been observed and were, until very recently, attributed exclusively to "measurement error"--instrumentation imprecision. A physicist at the University of California (Newman, 1982) noted that certain of these deviations appeared to be systematically associated with objects of intermediate size (several feet in diameter) at intermediate distances from one another (several feet to several miles). Systematic investigations of this once anomalous finding have served to corroborate the initial "hunch" such that significant theoretical revisions are now implied by regularities in current experimental data. To repeat the essential point of this illustrative example, a tendency to neglect or "look beyond" measurement error, although it may typically represent an appropriate and valuable heuristic device serving to reduce the cognitive demands of an inference task, may also reduce the probability that systematic fluctuations in data which have theoretical significance will be detected.

The implications of the results regarding measurement error for clinical diagnosis and treatment seem to have particular significance in the domain of psychometric measurement. An extensive body of empirical literature suggests that all psychological tests have limitations with regard to various types of reliability and validity, the majority of which are a direct result of limitations in the precision with which the characteristics under examination can be measured. It is conceivable that a cognitive predisposition toward neglect of measurement error may cause practicing clinicians to treat psychological assessment data as though it were wholly diagnostic when in fact it may frequently be only weakly correlated with the phenomenon which it purports to measure. It is interesting to note in this regard that those assessment instruments which systematic research suggests are the most dubious with regard to reliability and validity considerations (Fisher & Fisher, 1950; Holtzberg & Wexler, 1950; Swensen, 1957) remain extremely popular among practicing clinicians. The literature on illusory correlation (Chapman, 1967; Chapman & Chapman, 1967, 1969; Golding & Rorer, 1962; Smedslund, 1963; Ward & Jenkens, 1965), consistent with the findings of the present investigation, attribute this sustained popularity to the fact that individuals have a basic cognitive tendency to neglect random error in data in order to selectively perceive hypothesized associations between test responses and clinical syndromes which are in no way indicated in the assessment data.

Probe Check Likelihood Following Disconfirmatory, Versus Confirmatory, Outcomes

Recall that it was hypothesized that a greater proportion of probe checks would occur following disconfirmatory, than following confirmatory, outcomes due to the speculation that experimental methodology comes under greater scrutiny following empirical refutation of a proposed hypothesis. While a trend toward support of this hypothesis emerged in the data, the strength of this trend wasn't sufficient to reach statistical significance. The weak magnitude of this result should be considered, however, in light of the fact that the measurement designed to elicit the effect was a relatively insensitive one. Each subject was given only three opportunities to check on the status of the instrumentation during the task, and it is probably safe to assume that participants' emotional commitment to the hypothesis under test was relatively weak. That even a trend in the hypothesized direction emerged under these circumstances suggests that the basic theoretical question may be worthy of further investigation using a more sensitive methodology

(I'm questioning the veracity of one of my auxiliaries in an effort to avoid abandonment of the focal hypothesis!).

One final topic should be discussed in relation to the manner in which this question was addressed in the present investigation. Note that confirmatory and disconfirmatory outcomes had equal informational value in terms of the subject's focal hypothesis. Theoretical writings which have discussed the appropriateness of abandoning (versus modifying) disconfirmed hypotheses have argued that it is often more rational and appropriate to question one's auxiliary assumptions in the face of disconfirmation than in the face of confirmation due to the fact that confirmatory outcomes frequently carry more informational value than disconfirmatory outcomes. This assertion is based on the fact that when a precise empirical prediction is made (e.g., $c = v/r$) and then borne out by empirical observation, it becomes foolish to examine one's experimental methodology since the empirical outcome would have been extremely unlikely to occur had the hypothesis under test (including any auxiliaries with which it had been conjoined) been false. A disconfirmatory outcome under such circumstances, on the other hand, would be highly ambiguous, potentially attributable to any number of methodological flaws in the design of the experiment such that the truth of the focal hypothesis would remain undetectable in a single empirical test. Such explanations for the greater scrutiny of auxiliary assumptions following empirical refutation of a focal hypothesis do not presume the existence of a basic cognitive bias toward confirmation, but rather suggest that such a tendency is a wholly rational response to the parameters of the task situation. While the above argument may have considerable merit under many circumstances, the design of the present investigation with respect to the informational value of confirmatory and disconfirmatory outcomes was such that it was possible to isolate a purely cognitive bias toward increased scrutiny of experimental methodology following empirical refutation, from the rational motivations outlined above.

Confidence Ratings

Subjects' confidence ratings suggest that the behavioral tendency to postpone rejection of a disconfirmed hypothesis when the data contain stochastic error, was paralleled by a subjective experience of uncertainty with regard to the interpretation of data in these two conditions. Interestingly, the introduction of both measurement and stochastic error resulted in a great deal of variance in participants' subjective response to the feedback data. Note that while most confidence ratings in Conditions 2 through 4 were somewhat affected by the inclusion of error in the data, five subjects appeared to be "overwhelmed" by this manipulation. That is, they all indicated that they were "extremely uncertain" that they had located the true moisture line by the end of the task. Visual comparison of these subjects' responses with their actual distance from the true location of the critical line at the end of the experiment does not suggest any clear association between confidence ratings and likelihood of problem solution. It is interesting to speculate whether there might be large individual differences in how people engaged in hypothesis testing respond to data error. If this were the case then the same empirical outcome might lead one experimenter to make an enormous speculative leap while another experimenter might perform 100 replications of the same experiment before accepting the initial outcome as legitimate. Which of these strategies would be most likely to facilitate the discovery process would very likely depend on the specific research question being addressed. In the present task, for example, those subjects in the Stochastic Error Only Condition who were willing to make substantial leaps of inference on the basis of a single datum (e.g., Subject 6) tended to discover the true location of the moisture line much more rapidly than did subjects who made extensive disconfirming observations prior to revision (e.g., Subject 2). Given the potential existence of large individual differences in "conservatism" with respect to hypothesis revision, empirical investigation of any regularities in the task situation which might suggest prescriptive norms that should guide the choice of one strategy over another, would be extremely productive.

Proximity to True Moisture Line at Completion of Task

The data derived from measurement of the distance between participants' final placement of the hypothetical moisture line and the true location of the critical line tended to corroborate the inference that there exist large individual differences in the amount of data which must be accumulated before modification of a hypothesis is undertaken. Possible implications of this pattern in the data have already been discussed.

Additional Tribbles Requested

Because so much variance emerged in subjects' responses when they were asked how many additional tribbles they would need to plant in order to be absolutely certain that they had located the actual critical moisture line, it was difficult to draw any conclusive (or even speculative) inferences from this data. Unfortunately, based on subjects' frequent requests to have the phrase "absolutely certain" defined operationally, it is suspected that subjects' interpretations of the meaning of this question may have been highly idiosyncratic, making this an insensitive dependent measure.

Memory Test

Recall that subjects whose feedback data contained stochastic error tended to overestimate their predictive accuracy when asked how many of their experiments had come out as anticipated, while subjects in the other two groups tended either to accurately recall, or to slightly underestimate their predictive accuracy. The actual number of accurate experimental predictions did not differ across the four conditions. What did differ for subjects in Conditions 3 and 4 (Stochastic Error Only and both Measurement and Stochastic Error), however, was the amount of subjectively experienced ambiguity in the feedback data received throughout the experimental task (as suggested by results from both qualitative and quantitative dependent variables). A recent body of literature on attributions of causality and personal responsibility for outcomes (Ebbesen & Zeiss, 1976; Ross & Sicoly, 1979) suggests that biases in recall exist which cause one's own contributions to a joint product to be more easily and frequently recalled than are the contributions of others. A related phenomenon may account for the distortions in memory observed in the present investigation. That is, a general tendency toward egocentric attributional biases existing in all participants may have become manifest only in the context of ambiguity created by the introduction of stochastic error.

Directions for Future Research

Parametric Manipulations. Several possibilities for additional research are suggested by results of the present investigation. First, a parametric manipulation of both measurement and stochastic error would facilitate determination of whether some critical level exists for each of these independent variables, above which an impact on the dependent measures begins to emerge. (Note that the absence of any effect for measurement error may be due to the fact that the single level of this independent variable included in the present study didn't exceed this critical value.) Such a manipulation would also detect any systematic association between levels of independent variables and their magnitudes of effect on dependent measures. It seems plausible to suppose, for example, that if the amount of stochastic error contained in feedback data were increased, frequency of hypothesis revision on the part of subjects would decrease correspondingly. But would this decrease necessarily be linear? A number of interesting speculations about the effect of stochastic error on the inference process would be suggested by the finding that this association is best represented by a positively or negatively accelerating curve, for example.

Manipulation of "Theoretical" Auxiliaries. A second direction for future research has been alluded to throughout the dissertation. The manipulation of independent variables needs to be extended out of the domain of "instrumentational" auxiliaries into the domain of "theoretical" auxiliaries in order to provide a more comprehensive picture of the impact of auxiliary hypotheses on the inference process.

Which Will Be Questioned in the Face of Disconfirmation: Hypothesis or Auxiliaries? A third interesting question not addressed in the present investigation is whether there may exist parameters of the problem situation, other than the presence of data error, which influence the likelihood that a disconfirmatory outcome will lead to hypothesis revision rather than examination of auxiliaries. Such parameters might include: (1) The degree of prior commitment to the hypothesis under test; (2) The relative weight assigned to a given empirical datum as determined by factors other than those associated with measurement and/or stochastic error; and (3) The number of competing hypotheses which account equally well for an observed experimental outcome.

Further Investigation of the Impact of Data Error on the Inference Process. Systematic study of the effect of data error on the hypothesis-testing process is a new domain of inquiry within the psychology of science which deserves further investigation in its own right. Any number of paradigms could be employed in which subjects would be asked to detect lawful patterns in a universe where experimenter inputs were causally related to empirical outcomes, but in a probabilistic or imprecise fashion.

Summary and Conclusions

The results of the current study indicate that people do not invariably (or almost invariably) perseverate in testing falsified hypotheses, as earlier empirical work in the psychology of science has tended to suggest. Rather, the likelihood that a disconfirmed hypothesis will be revised in order to incorporate anomalous data appears to be strongly influenced by the nature of the falsifying data received by the investigator. When a certain proportion of the data is known, a priori, to have been generated at random, the researcher may manifest a strong tendency to replicate disconfirmatory outcomes many times before coming to accept them as conclusive. On the other hand, when the data is known to be free of such (stochastic) error, or to have some amount of random measurement error associated with each datum received, the typical investigator will readily revise a disconfirmed hypothesis, frequently on the basis of a single experiment.

In general, the present study suggests that attempts to formulate universal "rules" regarding the appropriateness of confirmatory versus disconfirmatory reasoning strategies, probably hold little promise for uncovering the cognitive mechanisms that underlie (or should underlie) scientific thinking. Instead, investigations of the unique parameters of the inference task which define the efficacy of one reasoning strategy over the other, seem more likely to lead ultimately to the design of prescriptions for better scientific inquiry.

APPENDIX A

Instructions to Subjects in the No Error, Measurement Error Only, and Stochastic Error Only Conditions, Followed by Questionnaire Completed by Subjects at End of Task

Instructions: No Error Condition (1)

You are a scientist, investigating an unexplored planet, Ethereus. Right now, you are orbiting Ethereus in a spaceship. From your spaceship, you can conduct a variety of controlled experiments. Previous research has shown that certain life forms exist on the planet, but the conditions which support these life forms are very poorly understood. Your research project will involve an attempt to determine what conditions promote the growth of a plant, the tribble, found in certain regions of Ethereus. The tribble was selected as the focus of this initial investigation because earlier work suggests that its survival depends only on the amount of moisture present in the soil. It is suspected that above a certain moisture content, tribbles grow. Below this moisture content, the plants die. Your task is to determine what this critical level is by systematically planting tribbles at various points on the planet's surface and seeing whether or not they survive at these locations. Each of the points you select for planting will correspond to a certain moisture level. The site of the investigation will be a 250,000 square-mile area encompassing a large portion of the planet's southern hemisphere. There is a display of this area on the computer console in front of you.

Fortunately, research has established that the distribution of moisture in the planet's soil is remarkably regular. The percentage of soil moisture on your viewing screen increases uniformly from west (left side of screen) to east (right side of screen). Preliminary data indicates that tribbles can survive only when the percentage of moisture in the soil equals or exceeds a certain level indicated by this dotted line *(E indicates)*. Your job is to conduct a more thorough investigation of the tentative hypothesis that tribbles can survive only at moisture levels equal to or exceeding the level displayed on the screen by planting them east or west of this line and observing whether they live or die.

Each time you're ready to plant a tribble, a small dot will appear on the screen. This dot *(E indicates to S)* represents the planting location of your first tribble. It can be moved to the east or west of where it first appears very easily. If you want to move the dot to the east, just hold down the "repeat" (REPT) key and press the arrow beneath it on the right (→). To move the dot in a western direction, hold down the repeat key and press the arrow in the left (←).

You have resources available to plant a total of 8 tribbles.

In order that I can better understand what you're thinking about as you perform the task, I'd like you to go through the following steps each time you plant a tribble:

1. When the dot appears, position it over the location at which you would like to plant. After the planting area has been selected, press the "F" key for FINISHED. *(E demonstrates)*

No Error (p. 2)

2. At the bottom of the screen it now says, "PREDICT RESULTS; D = DIE, L = LIVE." Based on the data you have available to you, I'd like you to try and predict whether the tribble you're about to plant will live or die. Initially, you may feel you don't have enough information to warrant a reasonable prediction, but I'd like you to do your best.

3. As soon as you've formulated your prediction I'd like you to tell me, as best you can, the reason for your prediction.

4. Release the tribble by pressing the "L" or the "D" key. The tribble will then descend to the surface of the planet and either grow or die. The result will be radioed back to your space craft and displayed on your screen as either green or purple. *(E demonstrates)*

5. At this point, the lower portion of the computer screen says, "DO YOU WANT TO CHANGE THE HYPOTHETICAL LINE? TYPE Y OR N." What I'm interested in finding out is whether or not the data you've generated has led you to reject the location of the original critical moisture line in favor of a new location. I want you to relocate this line only when your data indicates that the line's present location is wrong and does not represent the actual critical moisture level. Do this by typing "Y" for YES, and repositioning the hypothesis line in the same manner as repositioning a tribble (by using the arrow keys and the repeat key). *(E demonstrates)* When you've reached a point that represents your new working hypothesis about where the critical moisture line should be, type "F" for FINISHED and the computer will give you your next tribble. If you don't want to move the line, type "N" for NO and the computer will make your next tribble available for planting. This procedure will be followed for each tribble you plant until you've planted a total of 8 tribbles. Do you have any questions?

(E clarifies any areas on which the subject remains unclear at this point.)

It would be helpful to me if you could "talk out loud" as you perform the task, sharing with me any inferences you may be making or reasoning strategies you may be using as you work. I'll be making a tape recording of any thoughts or perceptions you share, so that I can go back later and correlate these with specific responses you've made on the task.

Instructions: Measurement Error Only Condition (2)

You are a scientist, investigating an unexplored planet, Ethereus. Right now, you are orbiting Ethereus in a spaceship. From your spaceship, you can conduct a variety of controlled experiments. Previous research has shown that certain life forms exist on the planet, but the conditions which support these life forms are very poorly understood. Your research project will involve an attempt to determine what conditions promote the growth of a plant, the tribble, found in certain regions of Ethereus. The tribble was selected as the focus of this initial investigation because earlier work suggests that its survival depends only on the amount of moisture present in the soil. It is suspected that above a certain moisture content, tribbles grow. Below this moisture content, the plants die. Your task is to determine what this critical level is by systematically planting tribbles at various points on the planet's surface and seeing whether or not they survive at these locations. Each of the points you select for planting will correspond to a certain moisture level. The site of the investigation will be a 250,000 square-mile area encompassing a large portion of the planet's southern hemisphere. There is a display of this area on the computer console in front of you.

Fortunately, research has established that the distribution of moisture in the planet's soil is remarkably regular. The percentage of soil moisture on your viewing screen increases uniformly from west (left side of screen) to east (right side of screen). Preliminary data indicates that tribbles can survive only when the percentage of moisture in the soil equals or exceeds a certain level indicated by this dotted line *(E indicates)*. Your job is to conduct a more thorough investigation of the tentative hypothesis that tribbles can survive only at moisture levels equal to or exceeding the level displayed on the screen by planting them east or west of this line and observing whether they live or die.

Each time you're ready to plant a tribble, a short horizontal line will appear on the screen. This line *(E indicates to S)* represents an area in which you can plant your tribble. This area can be moved to the east or west of where it first appears very easily. If you want to move the area to the east, just hold down the "repeat" (REPT) key and press the arrow beneath it on the right (→). To move the area in a western direction, hold down the repeat key and press the arrow on the left (←).

Why is the planting area represented by a horizontal line instead of a dot? Because you are orbiting the planet from a distance of 500 miles, you are not able to plant tribbles at a precise location. Rather, tribbles will land within 15 miles of the location at which you aim. Although the tribble may land somewhat east, west, north, or south of the location you specify, only east-west error is of interest with respect to your moisture hypothesis, so this is the error you see represented on the screen by the horizontal line. By moving the line east or west on each trial you can specify a planting location. When the tribble is released, you will know that it has landed somewhere within the east-west range represented by the horizontal line, but you can be no more specific in your observations than this.

Measurement Error Only (p. 2)

You have resources available to plant a total of 8 tribbles.

In order that I can better understand what you're thinking about as you perform the task, I'd like you to go through the following steps each time you plant a tribble:

1. When the line appears, position it over the location at which you would like to plant. After the planting area has been selected, press the "F" key for FINISHED. *(E demonstrates)*

2. At the bottom of the screen it now says, "PREDICT RESULTS; D = DIE, L = LIVE." Based on the data you have available to you, I'd like you to try and predict whether the tribble you're about to plant will live or die. Initially, you may feel you don't have enough information to warrant a reasonable prediction, but I'd like you to do your best.

3. As soon as you've formulated your prediction I'd like you to tell me, as best you can, the reason for your prediction.

4. Release the tribble by pressing the "L" or the "D" key. The tribble will then descend to the surface of the planet and either grow or die. The result will be radioed back to your space craft and displayed on your screen as either green or purple. *(E demonstrates) (E shows index card 1 to S and inquires)*

5. At this point, the lower portion of the computer screen says, "DO YOU WANT TO CHANGE THE HYPOTHETICAL LINE? TYPE Y OR N." What I'm interested in finding out is whether or not the data you've generated has led you to reject the location of the original critical moisture line in favor of a new location. I want you to relocate this line only when your data indicates that the line's present location is wrong and does not represent the actual critical moisture level. Do this by typing "Y" for YES, and repositioning the hypothesis line in the same manner as repositioning a tribble (by using the arrow keys and the repeat key). *(E demonstrates)* When you've reached a point that represents your new working hypothesis about where the critical moisture line should be, type "F" for FINISHED and the computer will give you your next tribble. If you don't want to move the line, type "N" for NO and the computer will make your next tribble available for planting. This procedure will be followed for each tribble you plant until you've planted a total of 8 tribbles. Do you have any questions?

(E clarifies any areas on which the subject remains unclear at this point.)

It would be helpful to me if you could "talk out loud" as you perform the task, sharing with me any inferences you may be making or reasoning strategies you may be using as you work. I'll be making a tape recording of any thoughts or perceptions you share, so that I can go back later and correlate these with specific responses you've made on the task.

Instructions: Stochastic Error Only Condition (3)

You are a scientist, investigating an unexplored planet, Ethereus. Right now, you are orbiting Ethereus in a spaceship. From your spaceship, you can conduct a variety of controlled experiments. Previous research has shown that certain life forms exist on the plaent, but the conditions which support these life forms are very poorly understood. Your research project will involve an attempt to determine what conditions promote the growth of a plant, the tribble, found in certain regions of Ethereus. The tribble was selected as the focus of this initial investigation because earlier work suggests that its survival depends only on the amount of moisture present in the soil. It is suspected that above a certain moisture content, tribbles grow. Below this moisture content, the plants die. Your task is to determine what this critical level is by systematically planting tribbles at various points on the planet's surface and seeing whether or not they survive at these locations. Each of the points you select for planting will correspond to a certain moisture level. The site of the investigation will be a 250,000 square-mile area encompassing a large portion of the planet's southern hemisphere. There is a display of this area on the computer console in front of you.

Fortunately, research has established that the distribution of moisture in the planet's soil is remarkably regular. The percentage of soil moisture on your viewing screen increases uniformly from west (left side of screen) to east (right side of screen). Preliminary data indicates that tribbles can survive only when the percentage of moisture in the soil equals or exceeds a certain level indicated by this dotted line *(E indicates)*. Your job is to conduct a more thorough investigation of the tentative hypothesis that tribbles can survive only at moisture levels equal to or exceeding the level displayed on the screen by planting them east or west of this line and observing whether they live or die.

Each time you're ready to plant a tribble, a small dot will appear on the screen. This dot *(E indicates to S)* represents the planting location of your first tribble. It can be moved to the east or west of where it first appears very easily. If you want to move the dot to the east, just hold down the "repeat" (REPT) key and press the arrow beneath it on the right (→). To move the dot in a western direction, hold down the repeat key and press the arrow on the left (←).

You have resources available to plant a total of 8 tribbles.

Because you are orbiting the planet from a considerable distance, there are several layers of instrumentation involved in collecting data regarding whether or not your tribble survived. Specifically, your spaceship is equipped with a series of probes, one of which is sent down with every tribble you plant. Occasionally, the probe malfunctions and sends back random information totally unrelated to the true condition of the tribble. That is, what the probe says about the tribble bears no relation to its true condition. *(E explains)* Previous work involving such probes indicates that the rate of malfunction is approximately 25%, or one in four. Because this is an average, however, your own rate of probe malfunction may be slightly greater or less than 25%. Fortunately, your spaceship is equipped with a device which can check on the functioning of the probe and tell you when you have received random data resulting from instrument malfunction. Because of the financial costs associated with probe checks, you will be allowed to employ it on only two occasions.

You have resources available to plant a total of 8 tribbles.

In order that I can better understand what you're thinking about as you perform the task, I'd like you to go through the following steps each time you plant a tribble:

1. When the dot appears, position it over the location at which you would like to plant. After the planting area has been selected, press the "F" key for FINISHED. *(E demonstrates)*

2. At the bottom of the screen it now says, "PREDICT RESULTS; D = DIE, L = LIVE." Based on the data you have available to you, I'd like you to try and predict whether the tribble you're about to plant will live or die. Initially, you may feel you don't have enough information to warrant a reasonable prediction, but I'd like you to do your best.

3. As soon as you've formulated your prediction I'd like you to tell me, as best you can, the reason for your prediction, on the white sheet of paper in front of you.

4. Release the tribble by pressing the "L" or the "D" key. The tribble will then descend to the surface of the planet and either grow or die. The result will be radioed back to your space craft and displayed on your screen as either green or purple. *(E demonstrates)*

5. Next, decide whether or not you want to do a probe check on this trial. Indicate "Y" for YES or "N" for NO in accordance with your decision. *(E demonstrates)* Remember, you have only two probe checks available.

6. At this point, the lower portion of the computer screen says, "DO YOU WANT TO CHANGE THE HYPOTHETICAL LINE? TYPE Y OR N." What I'm interested in finding out is whether or not the data you've generated has led you to reject the location of the original critical moisture line in favor of a new location. I want you to relocate this line only when your data indicates that the line's present location is wrong and does not represent the actual critical moisture level. Do this by typing "Y" for YES, and repositioning the hypothesis line in the same manner as repositioning a tribble (by using the arrow keys and the repeat key). *(E demonstrates)* When you've reached a point that represents your new working hypothesis about where the critical moisture

line should be, type "F" for FINISHED and the computer will give you your next tribble. If you don't want to move the line, type "N" for NO and the computer will make your next tribble available for planting. This procedure will be followed for each tribble you plant until you've planted a total of 8 tribbles. Do you have any questions?

(E clarifies any areas on which the subject remains unclear at this point.)

It would be helpful to me if you could "talk out loud" as you perform the task, sharing with me any inferences you may be making or reasoning strategies you may be using as you work. I'll be making a tape recording of any thoughts or perceptions you share, so that I can go back later and correlate these with specific responses you've made on the task.

Date ________________ Condition: __________ Subject ID ___________

How certain are you that the true line is within 1/2 inch of where your present line is? (Circle number)

1	2	3	4	5	6	7
Extremely Uncertain		Somewhat Uncertain		Somewhat Certain		Extremely Certain

If you had an infinite supply of tribbles, how many would you need to plant in order to be absolutely <u>certain</u> of the true location of the line?

____

Approximately how many of your "experiments" came out as predicted, that is, you predicted "live" and the tribble lived, or you predicted "die" and the tribble died? ___________

I'd like you to try and describe how "good" you felt the data you received was. That is, try to describe how diagnostic/helpful/valid/differentiating/etc. it was. There is no clear or specific answer to this question. I'm just interested in your impression of the "quality" of the data.

____

____

____

____

____

____

____

APPENDIX B

Computer Program Used to Generate Experimental Task (Program Language: BASIC)

```
5  GOTO 90
10  IF PS = 1 THEN 20
12  HTAB 1: VTAB 1: GET A$: IF A$ = A$(AC) THEN
AC(AC) = AC(AC) + 1: RETURN
14 AC = AC + 1:A$(AC) = A$:AC(AC) = 1: RETURN

20  IF RC = 0 THEN AC = AC + 1:RC = AC(AC)
22 A$ = A$(AC):RC = RC - 1
24  IF  ASC (A$) < 22 THEN  RETURN
26  FOR W = 1 TO 812: NEXT W: HOME : HTAB 20:
VTAB 23: PRINT A$;: FOR W = 1 TO 812: NEXT W:
RETURN
90  LOMEM: 16384:D$ =  CHR$ (4):FT$ = "TSK-PRMS":FS$
= "SUB-DAT": DIM A$(500): DIM AC(500): DIM IT$(20):
DIM TA$(41): PRINT D$;"BLOADBINARY": POKE 232,0:
POKE 233,3: FOR N = 1 TO 25:AD$ = AD$ + "-":
NEXT N
100  REM PRIME-START
110  TEXT :SD = 1
120  HOME : HTAB 8: VTAB 12: PRINT "PLAY OR REPLAY?
P OR R ";: GET A$: IF (A$ <  > "P" AND A$ <
> "R") THEN 120
130  HOME : HTAB 5: VTAB 12: PRINT "INSERT TASK
DISK AND HIT RETURN ";: GET B$: HOME : HTAB 15:
VTAB 12: PRINT "DISK WAIT"
140  IF A$ = "P" THEN PS = 0: PRINT : PRINT D$"OPEN"FT$
: PRINT D$"READ"FT$: INPUT LE,SE,MT,MC,ET,EC,TD,HS,RX:
PRINT D$"CLOSE"FT$: PRINT D$:TM = MT:CM = MC:TE
= ET:CE = EC: GOTO 300
145 PS = 1: GOSUB 8100: GOSUB 8200: HOME : HTAB
4: VTAB 12: PRINT "DATA ONLY OR RECREATE? D OR
R ";: GET A$: IF A$ = "D" THEN TT = TM + TE +
TD: GOTO 2840
150  HOME : HTAB 12: VTAB 12: PRINT "SOUND? Y
OR N ";: GET A$: IF A$ = "N" THEN SD = 0
300 HR = 140 - LE - 2:TS =  INT (159 / (TM +
TE + 1)):TL = LE: IF LE = 0 THEN TL = 1
400  FOR N = 776 TO 944: POKE N,6: NEXT N: FOR
N = 0 TO TM + TE: POKE 776 + N * (TS - 2),146:
NEXT N: POKE 776 + (TM + 1) * (TS - 2),0
1000  REM TSK-START
1005 AC = 0:DC = TD:TT = 0: IF TD > 0 THEN  TEXT
: HOME : VTAB 12: HTAB 19: PRINT "DEMO": FOR
W = 1 TO 1200: NEXT W
1100  REM TSK-RESTART
1125 TC = 0:CC = 0:HX = HS: HGR : HCOLOR= 3:
```

```
SCALE= 1: DRAW 2 AT HX,0: DRAW 2 AT HX + 1,0
1130 BY = TS * (TM + 1) + 1: HCOLOR= 2: HPLOT
HX,0 TO 0,0 TO 0,BY TO HX,BY: HCOLOR= 1: HPLOT
HX + 1,0 TO 279,0 TO 279,BY TO HX + 1,BY
2000  REM TSK-DRIVER
2010  GOSUB 3000: GOSUB 4000: GOSUB 5000
2015 TA$(TT) = TA$
2020  IF PS = 1 THEN  HOME : VTAB 22: GOSUB 6000:
PRINT TA$;: GOSUB 9999
2229 DC = DC - 1: IF DC = 0 THEN  HGR : TEXT
: HOME : VTAB 12: HTAB 17: PRINT "RESTART": FOR
W = 1 TO 1200: NEXT W: GOTO 1100
2290  IF TC < TM THEN 2000
2300  REM EXTRA-TRIBBLES
2310  IF TE = 0 THEN 2400
2312  TEXT : HOME : HTAB 18: VTAB 12: PRINT "WAIT";:
FOR W = 1 TO 1200: NEXT W: HTAB 16: PRINT "YOU
HAVE": HTAB 20: VTAB 14: PRINT TE" EXTRA TRIBBLES";:
IF CE > 0 THEN  HTAB 20: VTAB 16: PRINT CE" EXTRA
CHECKS";
2314  HCOLOR= 0: HPLOT 0,BY TO 279,BY:BY = TS
* (TM + TE + 1) + 1: HCOLOR= 2: HPLOT 0,0 TO
0,BY TO HX,BY: HCOLOR= 1: HPLOT HX + 1,BY TO
279,BY TO 279,0
2315  FOR W = 1 TO 1200: NEXT W: POKE  - 16304,0:
POKE 776 + (TM + 1) * (TS - 2),146: POKE 776
+ (TM + TE + 1) * (TS - 2),0: HCOLOR= 3: SCALE=
1: DRAW 2 AT HX,0: DRAW 2 AT HX + 1,0:TM = TM
+ TE:CM = CE:TE = 0:CE = 0:CC = 0: GOTO 2290
2400  REM POST-COLLECT
2410  IF PS = 0 THEN  HOME : GOSUB 9999
2800  REM TSK-COMPLETE
2810  IF PS = 1 THEN 2840
2820  TEXT : HOME : HTAB 10: VTAB 10: PRINT "DO
YOU WANT TO SAVE": HTAB 10: VTAB 12: PRINT "THIS
SUBJECT'S DATA?": HTAB 17: VTAB 14: PRINT "Y
OR N";: GET A$: IF A$ = "Y" THEN  GOSUB 8100:
GOSUB 8300
2830  GOTO 100
2840  TEXT : HOME : GOSUB 6000
2845  FOR N = 1 TO TT: PRINT TA$(N): PRINT :
IF N / 10 =  INT (N / 10) THEN  GOSUB 9999: HOME
: GOSUB 6000
2850  NEXT N
2855 N = N - 1: IF N / 10 <  >  INT (N / 10)
THEN  GOSUB 9999
2860  GOTO 100
3000  REM  TRIBBLE
```

```
3010 TC = TC + 1:TT = TT + 1:TY = TC * TS + 1
3200  REM POSITION
3210  HOME : VTAB 22: HTAB 5: PRINT "TRIBBLE";:
HTAB 30: VTAB 24: PRINT "= FINISHED";: HTAB 28:
VTAB 24: FLASH : PRINT "F";: HTAB 30: VTAB 22:
PRINT "<-";: HTAB 33: PRINT "->";: NORMAL
3212  VTAB 22: HTAB 13: PRINT "NO. ";TC;
3214  POKE 946,70: POKE 947,0: POKE 966,205:
POKE 945,30
3218  GOSUB 9200:TX = LE + 1 +  INT ((277 - 2
* LE) * RN): HCOLOR= 3: HPLOT TX - TL,TY TO TX
+ TL,TY:TI$ =  RIGHT$ ( STR$ (TX + 1000),3)
3220  GOSUB 10: IF A$ <  >  CHR$ (8) THEN 3240
3232  IF TX > LE + 1 THEN  XDRAW 1 AT TX + TL,TY:
XDRAW 1 AT TX - TL - 1,TY:TX = TX - 1: IF SD
= 1 THEN  CALL 948
3240  IF A$ <  >  CHR$ (21) THEN 3250
3242  IF TX < 279 - LE - 1 THEN  XDRAW 1 AT TX
- TL,TY: XDRAW 1 AT TX + TL + 1,TY:TX = TX +
1: IF SD = 1 THEN  CALL 948
3250  IF A$ <  > "F" THEN 3220
3300  REM PREDICT
3310  HOME : HTAB 13: VTAB 22: PRINT "PREDICT
RESULTS";: VTAB 24: HTAB 2: FLASH : PRINT "D";:
HTAB 24: PRINT "L";: NORMAL : HTAB 4: PRINT "=
DIE (PURPLE)";: HTAB 26: PRINT "= LIVE (GREEN)";
3350  GOSUB 10: IF (A$ <  > "D" AND A$ <  > "L")
THEN 3350
3360 TP$ = A$
3400  REM DROP
3410  HCOLOR= 0: HPLOT TX - TL,TY TO TX + TL,TY:
HOME : POKE 946,60: POKE 947,0: POKE 966,205:
FOR MP = 5 TO 255 STEP 5:XX = TX - TL +  INT
(2 * TL *  RND (1))
3443  IF MP < 175 THEN  HCOLOR= 3: HPLOT XX,TY
TO XX + 1,TY
3444  IF SD = 1 THEN  POKE 945,MP: CALL 948
3446  HCOLOR= 0: HPLOT XX,TY TO XX + 1,TY: NEXT
MP
3500  REM LAND
3520  GOSUB 9200:LX = TX - LE +  INT ((2 * LE
+ 1) * RN):TL$ =  RIGHT$ ( STR$ (LX + 1000),3)
3600  REM GROW
3620 TV = 0:TV$ = "D": IF LX > RX THEN TV = 1:TV$
= "L"
3700  REM PROBE
3720  GOSUB 9200:PV = 0:PI$ = "B": IF SE < RN
THEN PV = 1:PI$ = "V"
```

```
3800  REM REPORT
3810  IF PV = 1 THEN TH = 2 - TV
3820  IF PV = 0 THEN  GOSUB 9200:TH = 1 +  INT
(2 * RN)
3830  HCOLOR= TH: HPLOT TX - TL,TY TO TX + TL,TY:TA$
= "D": IF TH = 1 THEN TA$ = "L"
3900  REM PROBE-CHECK
3910  IF CC = CM THEN :PC$ = "X": GOTO 3999
3920  HOME : HTAB 1: VTAB 22: PRINT "YOU HAVE
";CM - CC;" PROBE CHECKS LEFT";: HTAB 36: PRINT
"= YES";: HTAB 1: VTAB 24: PRINT "DO YOU WANT
TO CHECK THIS PROBE?";: HTAB 36: PRINT "= NO";:
FLASH : HTAB 34: PRINT "N";: HTAB 34: VTAB 22:
PRINT "Y";: NORMAL
3930  GOSUB 10: IF A$ = "N" THEN PC$ = A$: GOTO
3999
3942  IF A$ <  > "Y" THEN 3930
3943 CC = CC + 1: HOME : IF SD = 0 THEN 3950
3945  POKE 946,70: POKE 947,32: POKE 966,206:
FOR MP = 80 TO 250 STEP 5: POKE 945,MP: CALL
948: NEXT MP
3946  POKE 946,10: POKE 947,8: POKE 966,206:
FOR W = 1 TO 40: POKE 945,200 + 55 *  RND (1):
CALL 948: NEXT W
3947  POKE 946,70: POKE 947,16: POKE 966,238:
FOR MP = 201 TO 1 STEP  - 5: POKE 945,MP: CALL
948: NEXT MP
3950  HOME : HTAB 1: VTAB 22: PRINT "ORIGINAL
PROBE DATA WAS ";: FLASH : IF PV = 1 THEN  PRINT
"VALID";:PC$ = "V": GOTO 3990
3956  PRINT "UNRELIABLE";:PC$ = "B"
3990  NORMAL : FOR W = 1 TO 2400: NEXT W
3999 TF$ =  RIGHT$ ( STR$ (TX + 1000),3): RETURN

4000  REM HYPO
4004 HI$ =  RIGHT$ ( STR$ (HX + 1000),3)
4010  HOME : HTAB 1: VTAB 22: PRINT "DO YOU WANT
TO CHANGE";: HTAB 36: PRINT "= YES";: HTAB 1:
VTAB 24: PRINT "THE HYPOTHETICAL LINE?";: HTAB
36: PRINT "= NO";: HTAB 34: FLASH : PRINT "N";:
HTAB 34: VTAB 22: PRINT "Y";: NORMAL
4020  GOSUB 10:HM$ = A$: IF A$ = "N" THEN 4299
4030  IF A$ <  > "Y" THEN 4020
4040  SCALE= 1: POKE 946,20: POKE 947,4: POKE
966,206
4100  HOME : VTAB 23: HTAB 5: PRINT "HYPOTHESIS";:
HTAB 30: VTAB 24: PRINT "= FINISHED";: HTAB 28:
VTAB 24: FLASH : PRINT "F";: HTAB 30: VTAB 22:
```

```
PRINT "<-";: HTAB 33: PRINT "->";: NORMAL
4210  POKE 945,60: GOSUB 10: IF A$ < > CHR$
(8) THEN 4220
4215  IF HX < > 140 - HR THEN  XDRAW 2 AT HX
+ 1,0: XDRAW 2 AT HX - 1,0: HCOLOR= 1: HPLOT
HX,0: HPLOT HX,BY:HX = HX - 1: IF SD = 1 THEN
 CALL 948
4220  IF A$ < > CHR$ (21) THEN 4290
4225  IF HX < > 140 + HR THEN  XDRAW 2 AT HX,0:
XDRAW 2 AT HX + 2,0: HCOLOR= 2: HPLOT HX + 1,0:
HPLOT HX + 1,BY:HX = HX + 1: IF SD = 1 THEN
CALL 948
4290  IF A$ < > "F" THEN 4210
4299 HF$ =  RIGHT$ ( STR$ (HX + 1000),3): RETURN

5000 TN$ =  RIGHT$ ( STR$ (TC + 100),2): IF DC
> 0 THEN TN$ = "D" +  RIGHT$ (TN$,1)
5010 TA$ = TN$ + " " + TI$ + " " + TF$ + " "
+ TP$ + " " + TL$ + " " + TV$ + " " + PI$ + "
" + TA$ + " " + PC$ + " " + HM$ + " " + HI$ +
" " + HF$: RETURN
6000  PRINT "T   T   T  T  T  T P T P H  H
H": PRINT "N   I   F  P  L  V I A C M  I   F":
PRINT : RETURN
8100  REM INDEX
8105  HOME : HTAB 15: VTAB 12: PRINT "DISK WAIT"
8110  PRINT : PRINT D$"OPEN"FS$",L5000": PRINT
D$"READ"FS$",R0": FOR N = 1 TO 20: INPUT IT$(N):
NEXT N: PRINT D$"CLOSE"FS$: PRINT D$
8130  HOME : FOR N = 1 TO 20: HTAB 3 -  LEN (
STR$ (N)): PRINT N;" ";IT$(N): NEXT N
8150  HTAB 1: VTAB 23: PRINT "ENTER SUBJECT NUMBER
";: INPUT SN
8151  IF (SN < 1 OR SN > 20) THEN 8150
8152  IF (PS = 0 AND IT$(SN) < > AD$) THEN 8150
8154  IF (PS = 1 AND IT$(SN) = AD$) THEN 8150
8158  IF PS = 1 THEN 8199
8160  VTAB 23: HTAB 1: PRINT "TAG  "AD$;: HTAB
5: INPUT IT$(SN)
8165  IF  LEN (IT$(SN)) < 1 THEN 8160
8170 IT$(SN) =  LEFT$ (IT$(SN),25)
8199  RETURN
8200  REM GET-OLD-SUBJ
8210  HOME : HTAB 15: VTAB 12: PRINT "DISK WAIT"
8220  PRINT D$"OPEN"FS$",L5000": PRINT D$"READ"FS$",R";
SN: INPUT LE,SE,TM,CM,TE,CE,TD,HS,RX
8225  FOR N = 1 TO TM + TE + TD: INPUT TA$(N):
NEXT N
```

```
8230 N = 0
8240 N = N + 1: INPUT A$(N),AC(N)
8244  IF A$(N) = "RA" THEN A$(N) =  CHR$ (21)
8245  IF A$(N) = "LA" THEN A$(N) =  CHR$ (8)
8250  IF A$(N) <  > "END-OF-DATA" THEN 8240
8260  PRINT D$"CLOSE"FS$: PRINT D$: RETURN
8300  REM PUT-NEW-SUBJ
8310  HOME : HTAB 15: VTAB 12: PRINT "DISK WAIT"
8320  PRINT D$"OPEN"FS$",L5000": PRINT D$"WRITE"FS$",R0
": FOR N = 1 TO 20: PRINT IT$(N): NEXT N: PRINT
D$"WRITE"FS$",R";SN
8345  PRINT LE: PRINT SE: PRINT MT: PRINT MC:
PRINT ET: PRINT EC: PRINT TD: PRINT HS: PRINT
RX
8347  FOR N = 1 TO TD + MT + ET: PRINT TA$(N):
NEXT N
8350  FOR N = 1 TO AC
8352  IF A$(N) =  CHR$ (21) THEN A$(N) = "RA"
8354  IF A$(N) =  CHR$ (8) THEN A$(N) = "LA"
8356  IF ( ASC (A$(N)) > 64 AND  ASC (A$(N))
< 90 OR  ASC (A$(N)) = 46) THEN  PRINT A$(N):
PRINT AC(N)
8358  NEXT N
8360  PRINT "END-OF-DATA": PRINT 999999: PRINT
D$"CLOSE"FS$: PRINT D$: RETURN
9200  IF PS = 0 THEN A$(AC + 1) = "." +  RIGHT$
( STR$ (1E5 +  INT (1E5 *  RND (1))),5):AC(AC
+ 1) = 1
9250 AC = AC + 1:RN =  VAL (A$(AC)): RETURN
9999  HTAB 40: VTAB 24: GET B$: RETURN
10000  PR# 1: PRINT  CHR$ (9)"70N" CHR$ (9)"60R"
CHR$ (9)"15L" CHR$ (27)"E" CHR$ (27)"C" CHR$
(66) CHR$ (27)"N" CHR$ (18): LIST
```

APPENDIX C

Raw Data for All Experimental Subjects

DATA INTERPRETATION

Screen	Code	Item	Interpretation
XX	TN	Tribble Number	Number of last tribble. Restart after demo tribbles.
XXX	TI	Tribble Initial	Position where tribble is presented to subject.
XXX	TF	Tribble Final	Position to which subject moves tribble.
X	TP	Tribble Prediction	Subject's prediction. D = Die L = Live.
XXX	TL	Tribble Land	Position at which tribble actually lands.
X	TV	Tribble Vitality	Actual growth of tribble. D = Die L = Live.
X	PI	Probe Integrity	Probe condition. V = Valid B = Broken (Unreliable)
X	TA	Tribble Appearance	Info to subject from probe. D = Die L = Live.
X	PC	Probe Check	Used. N = NO V = yes/Valid B = yes/broken X = none left
X	HM	Hypo Move	Did supject opt to move line. Y = Yes N = No.
XXX	HI	Hypo Initial	Position of hypothetical line before tribble.
XXX	HF	Hypo Final	Position of Hypothetical line after tribble.

Box encloses what is shown on the screen.
X's will be numbers or letters.

Note. Hypo line at task begin = 110.
Real line = 210.
TV = L only when TL is <u>greater</u> than 210.

The letter B (for Broken) is used in PI and PC instead of U (for Unreliable) because U's are difficult to distinguish from V's (for Valid) on the television screen.

+ or +++ = live tribble; - or —— = dead tribble

CONDITION: 1 - NO ERROR

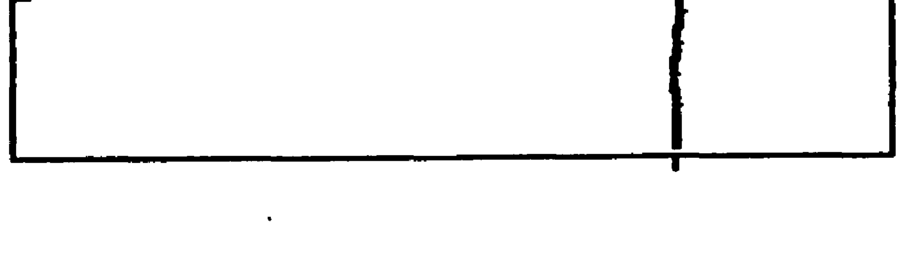
SUBJECT NUM: 1

SUBJECT TAG: 9 PHYSICS EXP 5TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	060	D	060	D	V	D	X	Y	110	111
01	115	122	L	122	D	V	D	X	Y	110	133
02	086	161	L	161	D	V	D	X	Y	133	182
03	275	206	L	206	D	V	D	X	Y	182	226
04	082	245	L	245	L	V	L	X	N	226	226
05	152	227	L	227	L	V	L	X	Y	226	215
06	238	219	L	219	L	V	L	X	Y	215	212
07	111	213	L	213	L	V	L	X	Y	212	209
08	068	210	L	210	D	V	D	X	Y	209	211
09	153	212	L	212	L	V	L	X	Y	211	210
10	029	211	L	211	L	V	L	X	Y	210	209
11	092	210	D	210	D	V	D	X	Y	209	210
12	234	211	L	211	L	V	L	X	N	210	210

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

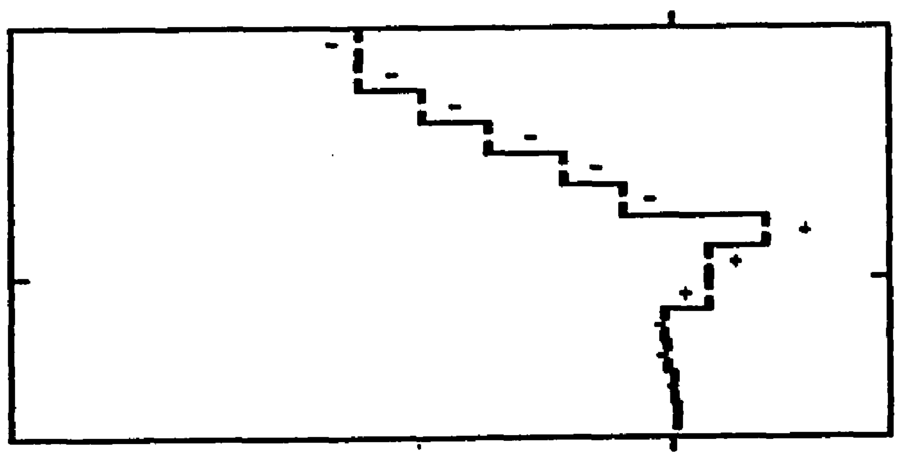
SUBJECT NUM: 2

SUBJECT TAG: 6 PHYSICS EXP 5TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	068	047	D	047	D	V	D	X	Y	110	110
01	028	102	D	102	D	V	D	X	N	110	110
02	222	121	L	121	D	V	D	X	Y	110	130
03	183	141	L	141	D	V	D	X	Y	130	151
04	256	165	L	165	D	V	D	X	Y	151	175
05	067	186	L	186	D	V	D	X	Y	175	194
06	207	203	L	203	D	V	D	X	Y	194	239
07	105	252	L	252	L	V	L	X	Y	239	221
08	046	230	L	230	L	V	L	X	N	221	221
09	253	214	D	214	L	V	L	X	Y	221	207
10	124	206	D	206	D	V	D	X	Y	207	208
11	069	207	D	207	D	V	D	X	Y	208	210
12	015	210	D	210	D	V	D	X	Y	210	211

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

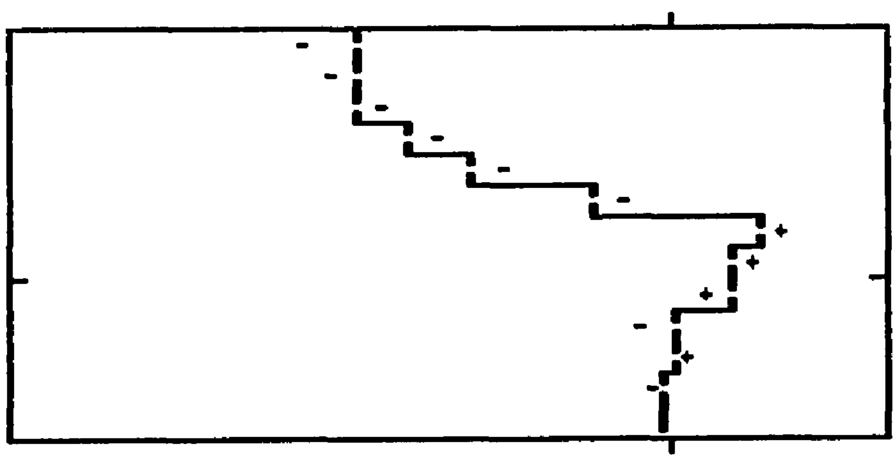
SUBJECT NUM: 3

SUBJECT TAG: 18 PHYSICS EXP 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	140	056	D	056	D	V	D	X	Y	110	110
01	169	093	L	093	D	V	D	X	N	110	110
02	001	102	L	102	D	V	D	X	N	110	110
03	106	118	L	118	D	V	D	X	Y	110	126
04	251	136	L	136	D	V	D	X	Y	126	146
05	151	157	L	157	D	V	D	X	Y	146	185
06	155	195	L	195	D	V	D	X	Y	185	238
07	217	245	L	245	L	V	L	X	Y	238	229
08	244	236	L	236	L	V	L	X	Y	229	229
09	119	221	L	221	L	V	L	X	Y	229	211
10	214	200	L	200	D	V	D	X	N	211	211
11	144	215	L	215	L	V	L	X	Y	211	207
12	072	204	L	204	D	V	D	X	N	207	207

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

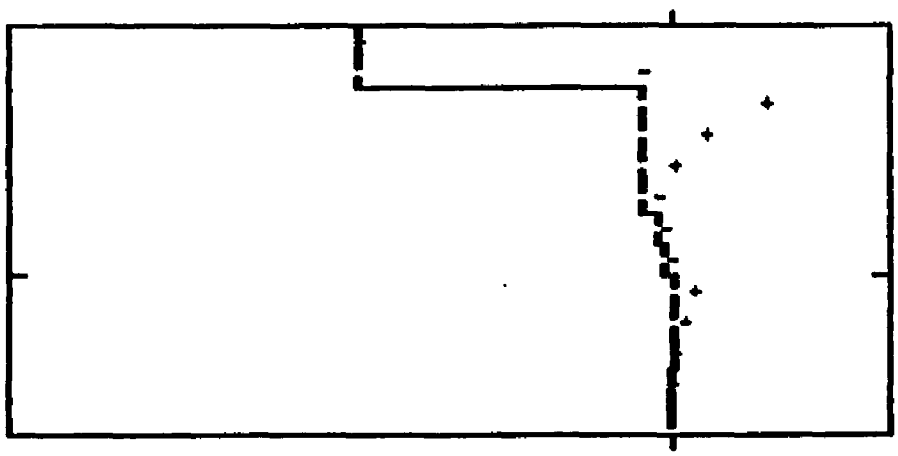
SUBJECT NUM: 4

SUBJECT TAG: 26 PHYSICS EXP 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	061	D	061	D	V	D	X	Y	110	109
01	115	111	L	111	D	V	D	X	N	110	110
02	086	201	L	201	D	V	D	X	Y	110	200
03	275	240	L	240	L	V	L	X	N	200	200
04	082	221	L	221	L	V	L	X	N	200	200
05	152	211	L	211	L	V	L	X	N	200	200
06	238	206	D	206	D	V	D	X	Y	200	205
07	111	208	D	208	D	V	D	X	Y	205	207
08	068	210	L	210	D	V	D	X	Y	207	210
09	153	217	L	217	L	V	L	X	N	210	210
10	029	214	L	214	L	V	L	X	N	210	210
11	092	211	L	211	L	V	L	X	Y	210	209
12	234	210	D	210	D	V	D	X	N	209	209

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

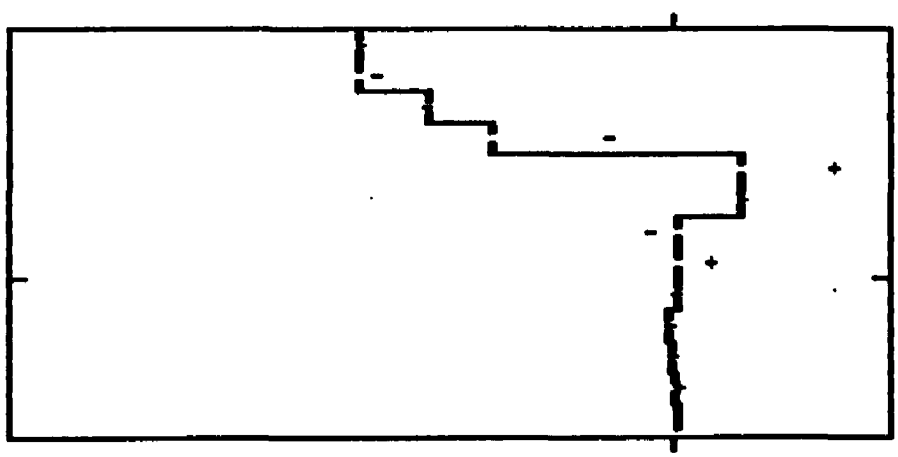
SUBJECT NUM: 5

SUBJECT TAG: 14 CHEMISTRY INORG 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	095	027	D	027	D	V	D	X	Y	110	110
01	185	111	L	111	D	V	D	X	N	110	110
02	047	116	L	116	D	V	D	X	Y	110	132
03	261	132	L	132	D	V	D	X	Y	132	152
04	030	190	L	190	D	V	D	X	Y	152	231
05	106	261	L	261	L	V	L	X	N	231	231
06	173	232	L	232	L	V	L	X	Y	231	211
07	016	203	D	203	D	V	D	X	N	211	211
08	265	222	L	222	L	V	L	X	N	211	211
09	045	211	L	211	L	V	L	X	Y	211	208
10	195	209	L	209	D	V	D	X	Y	208	209
11	057	210	L	210	D	V	D	X	Y	209	210
12	121	212	L	212	L	V	L	X	Y	210	211

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

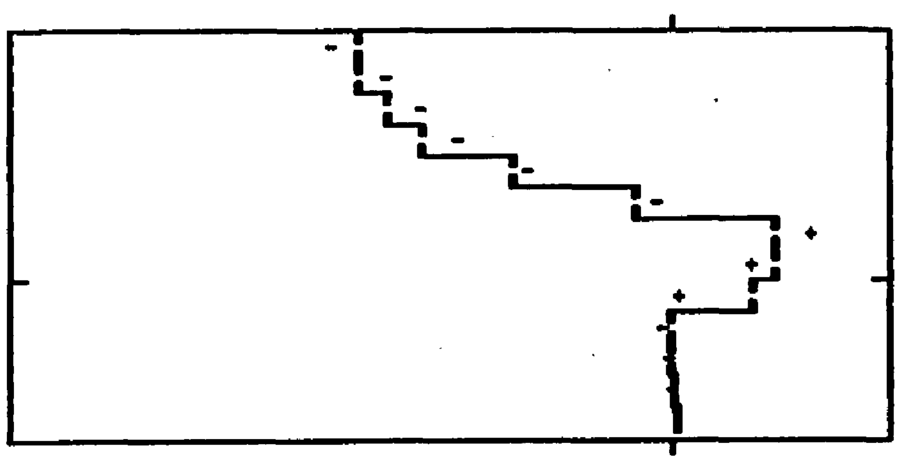
SUBJECT NUM: 6

SUBJECT TAG: 21 CHEMISTRY ORG 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	245	055	D	055	D	V	D	X	Y	110	104
01	230	102	D	102	D	V	D	X	N	110	110
02	276	119	L	119	D	V	D	X	Y	110	119
03	175	130	L	130	D	V	D	X	Y	119	130
04	097	142	L	142	D	V	D	X	Y	130	159
05	027	164	L	164	D	V	D	X	Y	159	198
06	115	205	L	205	D	V	D	X	Y	198	242
07	009	254	L	254	L	V	L	X	N	242	242
08	005	235	L	235	L	V	L	X	Y	242	235
09	007	212	D	212	L	V	L	X	Y	235	209
10	196	207	D	207	D	V	D	X	N	209	209
11	120	209	L	209	D	V	D	X	Y	209	210
12	117	210	L	210	D	V	D	X	Y	210	211

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

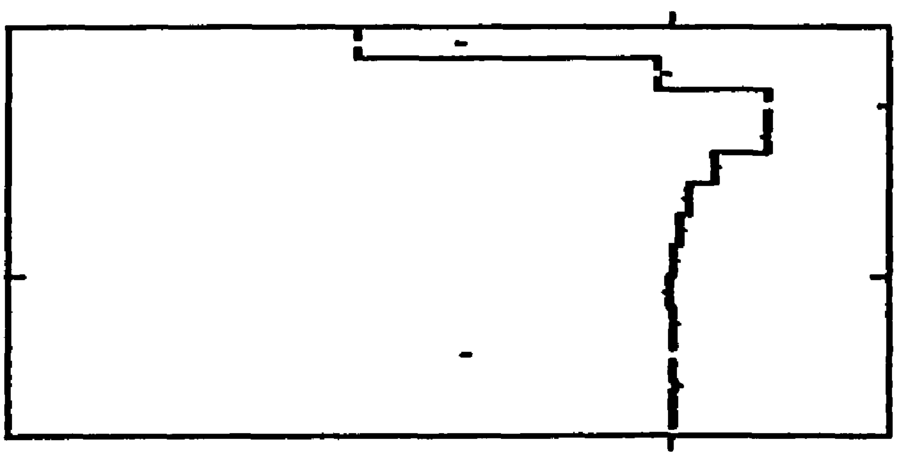
SUBJECT NUM: 7

SUBJECT TAG: 25 CHEMISTRY ORG 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	047	047	D	047	D	V	D	X	Y	110	108
01	043	143	L	143	D	V	D	X	Y	110	205
02	105	208	L	208	D	V	D	X	Y	205	240
03	109	277	L	277	L	V	L	X	N	240	240
04	122	240	L	240	L	V	L	X	Y	240	223
05	077	224	L	224	L	V	L	X	Y	223	215
06	093	215	L	215	L	V	L	X	Y	215	212
07	079	213	L	213	L	V	L	X	Y	212	210
08	223	211	L	211	L	V	L	X	Y	210	209
09	165	209	D	209	D	V	D	X	Y	209	210
10	167	211	L	211	L	V	L	X	N	210	210
11	145	145	D	145	D	V	D	X	N	210	210
12	212	212	L	212	L	V	L	X	N	210	210

PARAMETERS

LE	= 0
SE	= 0
TD	= 1
TS	= 8
CS	= 0
TE	= 4
CE	= 0
HS	= 110
RL	= 210



CONDITION: 1 - NO ERROR

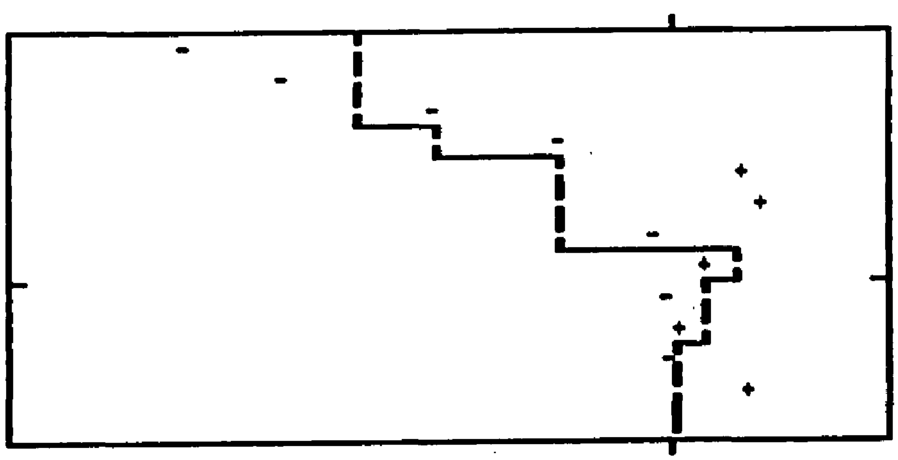
SUBJECT NUM: 8

SUBJECT TAG: 30 BIOLOGY MICRO 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	059	D	059	D	V	D	X	Y	110	108
01	115	055	D	055	D	V	D	X	N	110	110
02	086	086	L	086	D	V	D	X	N	110	110
03	275	134	L	134	D	V	D	X	Y	110	135
04	082	174	L	174	D	V	D	X	Y	135	174
05	152	232	L	232	L	V	L	X	N	174	174
06	238	238	L	238	L	V	L	X	N	174	174
07	111	204	L	204	D	V	D	X	Y	174	230
08	068	220	D	220	L	V	L	X	Y	230	220
09	153	208	D	208	D	V	D	X	N	220	220
10	029	212	D	212	L	V	L	X	Y	220	211
11	092	209	D	209	D	V	D	X	N	211	211
12	234	234	L	234	L	V	L	X	N	211	211

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

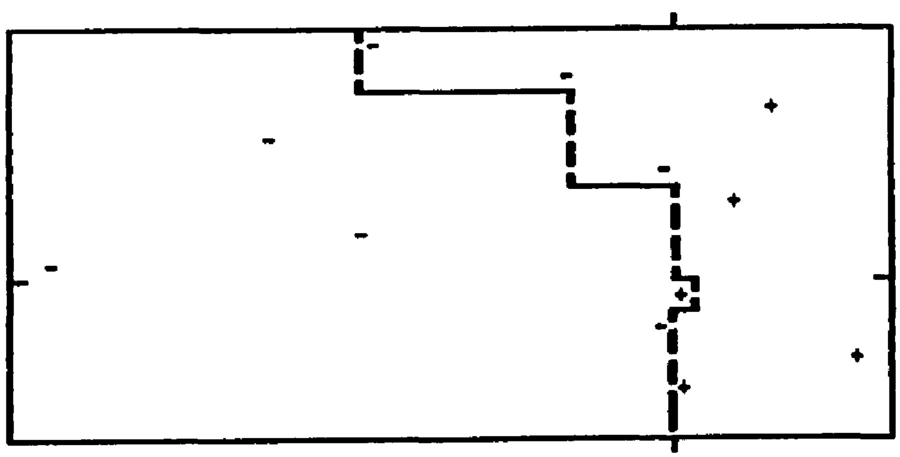
SUBJECT NUM: 9

SUBJECT TAG: 36 BIOLOGY BIOCHEM 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	069	D	069	D	V	D	X	Y	110	110
01	115	115	L	115	D	V	D	X	N	110	110
02	086	176	L	176	D	V	D	X	Y	110	177
03	275	241	L	241	L	V	L	X	N	177	177
04	082	082	D	082	D	V	D	X	N	177	177
05	152	207	L	207	D	V	D	X	Y	177	210
06	238	229	L	229	L	V	L	X	N	210	210
07	111	111	D	111	D	V	D	X	N	210	210
08	068	013	D	013	D	V	D	X	Y	210	216
09	153	212	D	212	L	V	L	X	Y	216	209
10	029	206	D	206	D	V	D	X	N	209	209
11	092	268	L	268	L	V	L	X	N	209	209
12	234	213	L	213	L	V	L	X	N	209	209

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 1 - NO ERROR

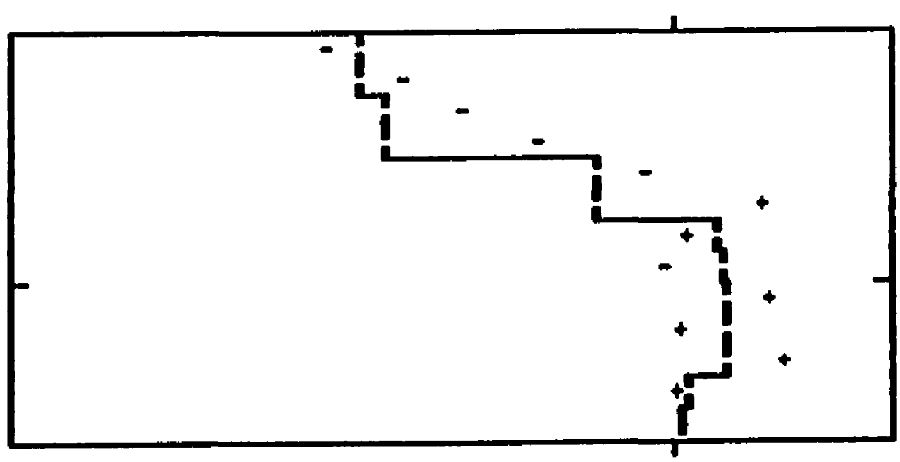
SUBJECT NUM: 10

SUBJECT TAG: 39 BIOLOGY MICRO 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	067	D	067	D	V	D	X	Y	110	107
01	115	100	D	100	D	V	D	X	N	110	110
02	086	124	L	124	D	V	D	X	Y	110	118
03	275	143	L	143	D	V	D	X	N	118	118
04	082	167	L	167	D	V	D	X	Y	118	185
05	152	201	L	201	D	V	D	X	N	185	185
06	238	238	L	238	L	V	L	X	Y	185	223
07	111	214	L	214	L	V	L	X	Y	223	225
08	068	207	L	207	D	V	D	X	Y	225	226
09	153	240	L	240	L	V	L	X	N	226	226
10	029	212	L	212	L	V	L	X	Y	226	226
11	092	245	D	245	L	V	L	X	Y	226	214
12	234	211	L	211	L	V	L	X	Y	214	212

PARAMETERS

LE = 0
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

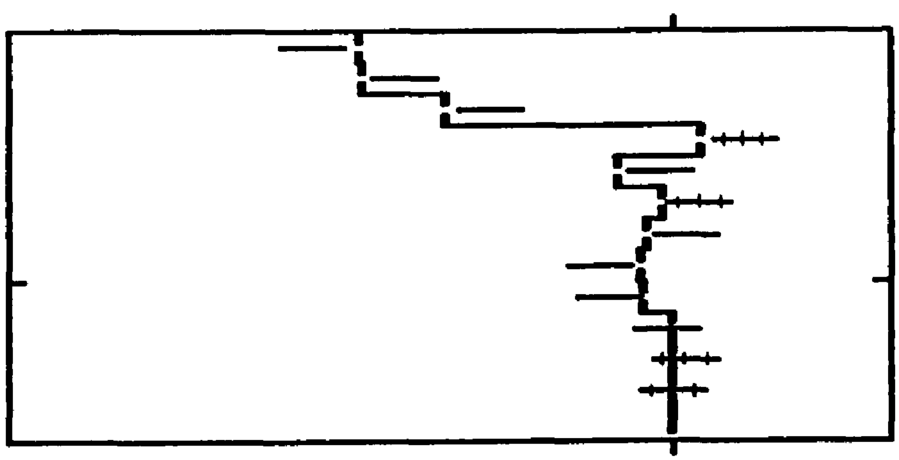
SUBJECT NUM: 1

SUBJECT TAG: 7 PHYSICS EXP 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	172	029	D	035	D	V	D	X	Y	110	110
01	137	096	D	103	D	V	D	X	Y	110	111
02	243	125	L	131	D	V	D	X	Y	111	137
03	101	152	L	161	D	V	D	X	Y	137	218
04	059	233	L	242	L	V	L	X	Y	218	192
05	072	206	L	208	D	V	D	X	Y	192	206
06	043	218	L	218	L	V	L	X	Y	206	201
07	179	214	L	221	L	V	L	X	Y	201	199
08	182	187	D	185	D	V	D	X	Y	199	200
09	134	190	D	194	D	V	D	X	Y	200	209
10	098	208	L	204	D	V	D	X	N	209	209
11	169	214	D	221	L	V	L	X	N	209	209
12	045	210	D	217	L	V	L	X	N	209	209

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

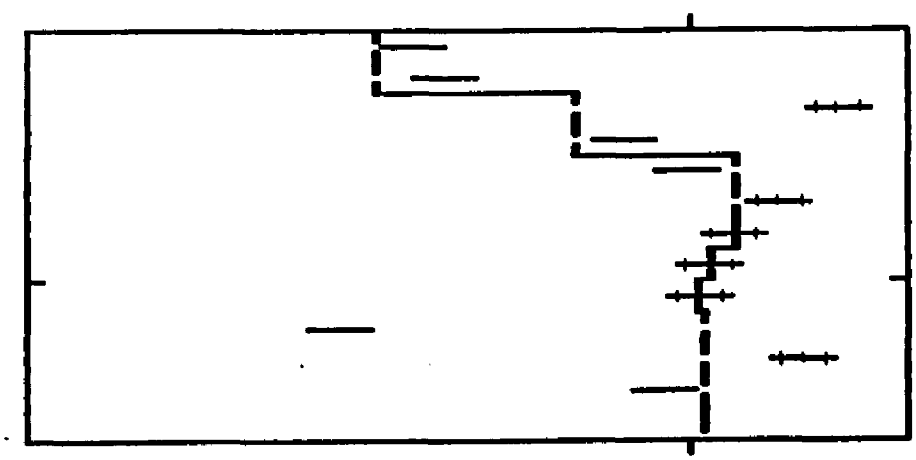
SUBJECT NUM: 2

SUBJECT TAG: 8 PHYSICS EXP 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	065	036	D	043	D	V	D	X	Y	110	110
01	059	122	L	132	D	V	D	X	N	110	110
02	180	132	L	140	D	V	D	X	Y	110	173
03	161	257	L	253	L	V	L	X	N	173	173
04	094	189	D	182	D	V	D	X	Y	173	224
05	034	209	D	206	D	V	D	X	N	224	224
06	122	238	L	230	L	V	L	X	N	224	224
07	111	224	L	219	L	V	L	X	Y	224	216
08	242	216	D	222	L	V	L	X	Y	216	212
09	098	213	L	208	D	V	D	X	Y	212	214
10	174	099	D	102	D	V	D	X	N	214	214
11	179	246	L	239	L	V	L	X	N	214	214
12	157	202	D	209	D	V	D	X	N	214	214

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

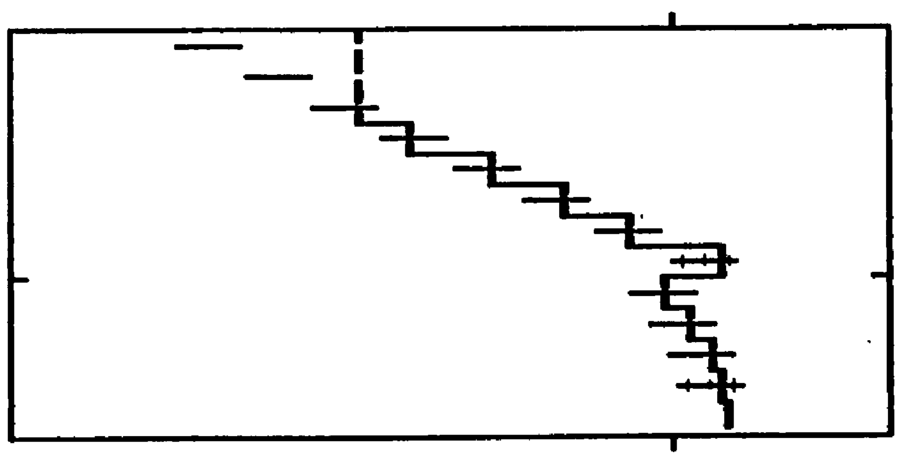
SUBJECT NUM: 3

SUBJECT TAG: 11 PHYSICS EXP 1ST YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	248	031	D	023	D	V	D	X	Y	110	109
01	133	063	D	061	D	V	D	X	N	110	110
02	126	085	D	093	D	V	D	X	N	110	110
03	013	106	L	113	D	V	D	X	Y	110	126
04	122	128	L	138	D	V	D	X	Y	126	152
05	222	151	L	154	D	V	D	X	Y	152	175
06	255	173	L	170	D	V	D	X	Y	175	196
07	019	196	L	187	D	V	D	X	Y	196	225
08	245	220	L	218	L	V	L	X	Y	225	207
09	089	207	L	197	D	V	D	X	Y	207	215
10	242	213	L	203	D	V	D	X	Y	215	222
11	037	219	L	209	D	V	D	X	Y	222	225
12	072	222	L	226	L	V	L	X	Y	225	227

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

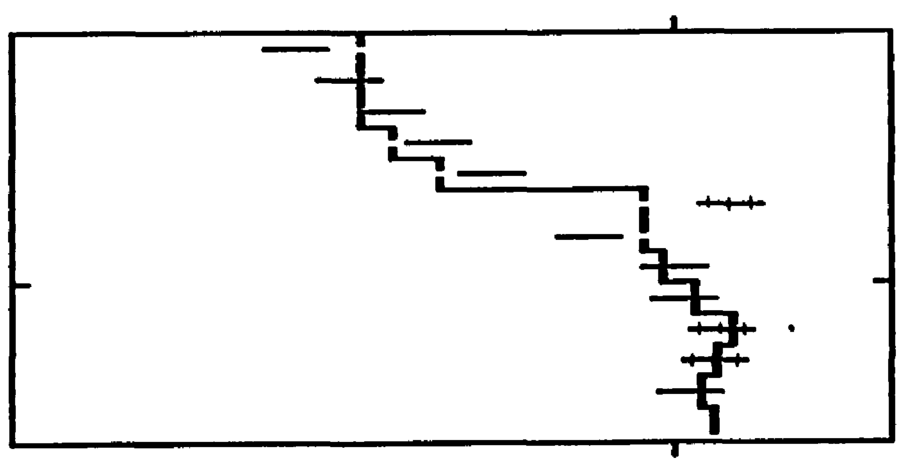
SUBJECT NUM: 4

SUBJECT TAG: 23 PHYSICS EXP 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	127	056	D	052	D	V	D	X	Y	110	108
01	116	090	D	080	D	V	D	X	N	110	110
02	188	107	D	116	D	V	D	X	N	110	110
03	018	120	L	123	D	V	D	X	Y	110	120
04	142	135	L	136	D	V	D	X	Y	120	135
05	044	152	L	155	D	V	D	X	Y	135	200
06	071	228	L	222	L	V	L	X	N	200	200
07	071	183	D	187	D	V	D	X	Y	200	206
08	250	210	L	203	D	V	D	X	Y	206	216
09	112	213	D	207	D	V	D	X	Y	216	228
10	159	225	L	220	L	V	L	X	Y	228	223
11	217	223	L	224	L	V	L	X	Y	223	218
12	190	215	L	207	D	V	D	X	Y	218	222

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

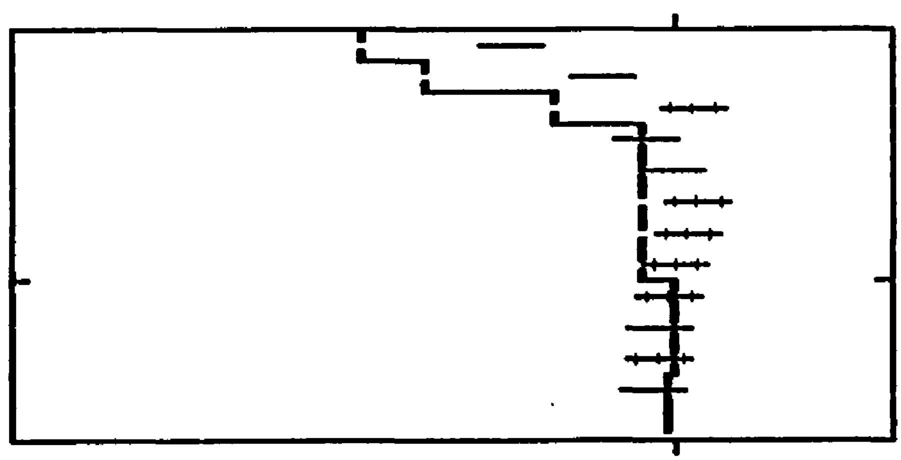
SUBJECT NUM: 5

SUBJECT TAG: 13 CHEMISTRY ORG 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	191	035	D	036	D	V	D	X	Y	110	110
01	215	158	L	148	D	V	D	X	Y	110	130
02	032	187	L	182	D	V	D	X	Y	130	171
03	216	216	L	226	L	V	L	X	Y	171	199
04	216	201	L	199	D	V	D	X	N	199	199
05	082	209	L	201	D	V	D	X	N	199	199
06	018	217	L	227	L	V	L	X	N	199	199
07	137	214	L	214	L	V	L	X	N	199	199
08	121	210	L	215	L	V	L	X	Y	199	209
09	239	208	L	212	L	V	L	X	N	209	209
10	153	205	L	200	D	V	D	X	N	209	209
11	036	205	D	212	L	V	L	X	Y	209	207
12	072	203	D	210	D	V	D	X	N	207	207

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

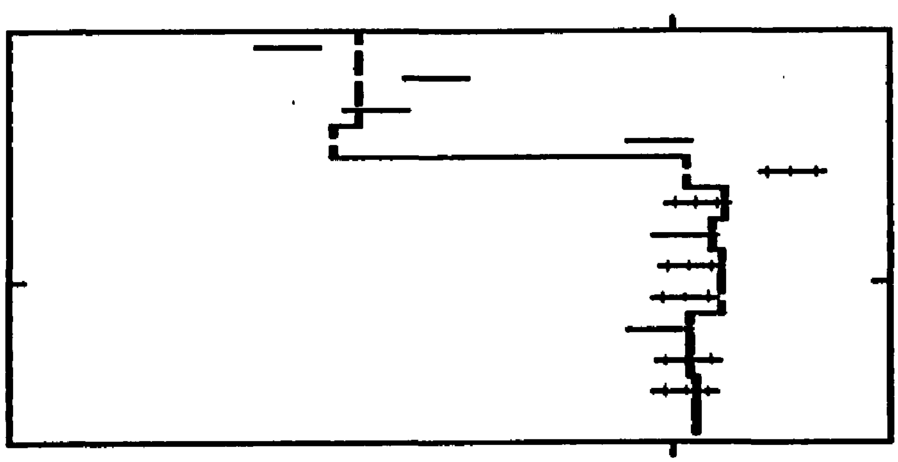
SUBJECT NUM: 6

SUBJECT TAG: 22 CHEMISTRY INORG 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	063	D	061	D	V	D	X	Y	110	112
01	117	088	L	090	D	V	D	X	N	110	110
02	090	135	L	130	D	V	D	X	N	110	110
03	265	116	L	108	D	V	D	X	Y	110	102
04	086	206	L	209	D	V	D	X	Y	102	214
05	151	248	L	245	L	V	L	X	Y	214	226
06	231	218	D	216	L	V	L	X	Y	226	222
07	113	214	D	206	D	V	D	X	Y	222	225
08	073	216	L	213	L	V	L	X	N	225	225
09	152	214	D	216	L	V	L	X	Y	225	215
10	037	206	D	208	D	V	D	X	N	215	215
11	096	215	L	225	L	V	L	X	Y	215	217
12	227	214	L	220	L	V	L	X	N	217	217

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

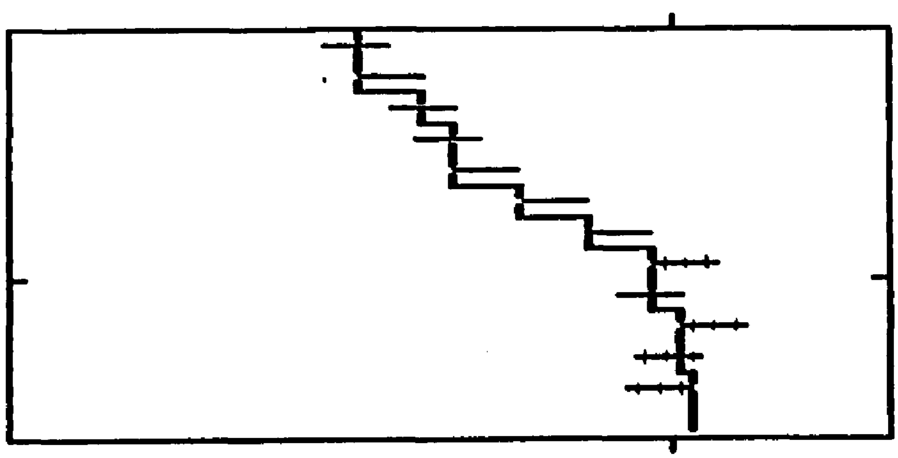
SUBJECT NUM: 7

SUBJECT TAG: 33 CHEMISTRY BIOCHEM 3 YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	070	D	068	D	V	D	X	Y	110	110
01	117	110	L	112	D	V	D	X	N	110	110
02	090	121	L	116	D	V	D	X	Y	110	130
03	265	131	L	123	D	V	D	X	Y	130	140
04	086	139	L	142	D	V	D	X	N	140	140
05	151	151	L	148	D	V	D	X	Y	140	161
06	231	173	L	171	D	V	D	X	Y	161	183
07	113	193	L	185	D	V	D	X	Y	183	203
08	073	214	L	211	L	V	L	X	N	203	203
09	152	203	L	205	D	V	D	X	Y	203	212
10	037	223	L	225	L	V	L	X	N	212	212
11	096	209	L	219	L	V	L	X	Y	212	216
12	227	206	L	212	L	V	L	X	N	216	216

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

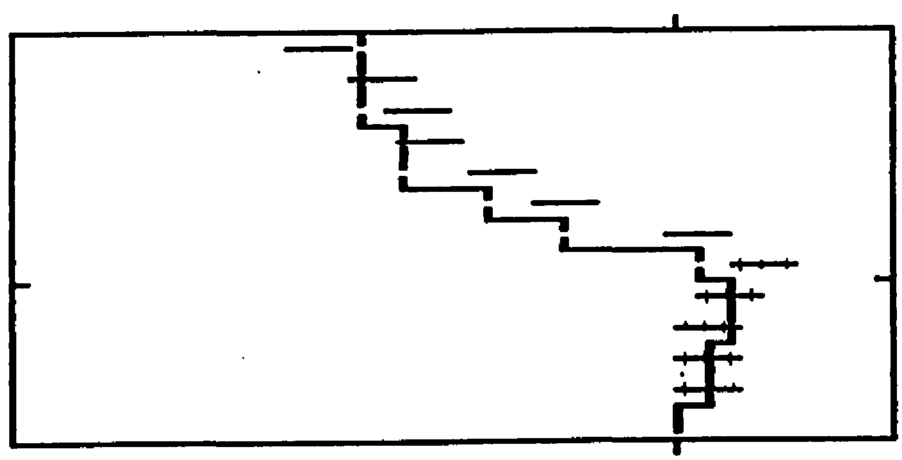
SUBJECT NUM: 8

SUBJECT TAG: 29 BIOLOGY MICRO 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	075	D	073	D	V	D	X	Y	110	103
01	117	097	L	099	D	V	D	X	N	110	110
02	090	117	L	112	D	V	D	X	N	110	110
03	265	128	L	120	D	V	D	X	Y	110	123
04	086	132	L	135	D	V	D	X	N	123	123
05	151	155	L	152	D	V	D	X	Y	123	150
06	231	175	L	173	D	V	D	X	Y	150	174
07	113	217	L	209	D	V	D	X	Y	174	217
08	073	238	L	235	L	V	L	X	Y	217	227
09	152	227	L	229	L	V	L	X	N	227	227
10	037	220	L	222	L	V	L	X	Y	227	220
11	096	220	L	230	L	V	L	X	N	220	220
12	227	220	L	226	L	V	L	X	Y	220	210

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

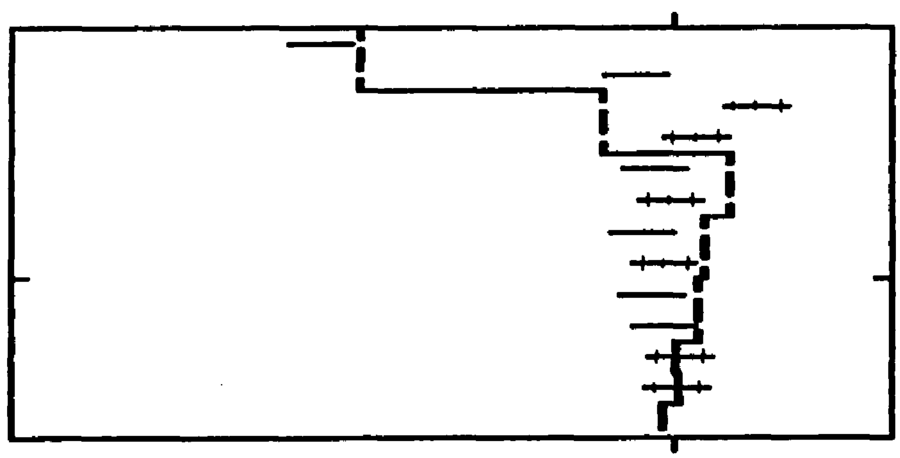
SUBJECT NUM: 9

SUBJECT TAG: 35 BIOLOGY MICRO 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	191	053	D	054	D	V	D	X	Y	110	106
01	215	098	D	088	D	V	D	X	N	110	110
02	032	198	L	193	D	V	D	X	Y	110	187
03	216	236	L	246	L	V	L	X	N	187	187
04	216	217	L	215	L	V	L	X	Y	187	227
05	082	204	L	196	D	V	D	X	N	227	227
06	018	209	L	219	L	V	L	X	Y	227	219
07	137	200	D	200	D	V	D	X	N	219	219
08	121	207	L	212	L	V	L	X	Y	219	217
09	239	203	L	207	D	V	D	X	N	217	217
10	153	207	L	202	D	V	D	X	Y	217	210
11	036	212	L	219	L	V	L	X	Y	210	211
12	072	211	D	218	L	V	L	X	Y	211	206

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 2 - MEASUREMENT ERROR

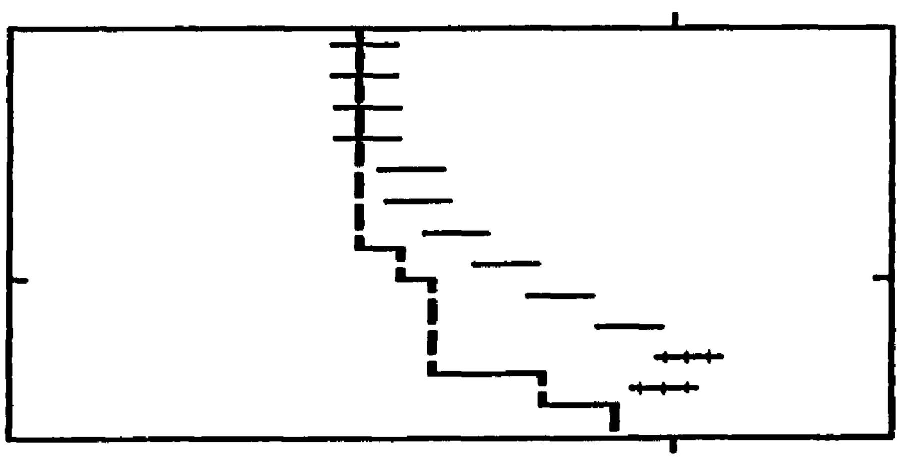
SUBJECT NUM: 10

SUBJECT TAG: 37 BIOLOGY BIOCHEM 5TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	066	D	064	D	V	D	X	Y	110	107
01	117	112	D	114	D	V	D	X	N	110	110
02	090	112	D	107	D	V	D	X	N	110	110
03	265	113	D	105	D	V	D	X	N	110	110
04	086	113	D	116	D	V	D	X	N	110	110
05	151	127	L	124	D	V	D	X	N	110	110
06	231	129	L	127	D	V	D	X	N	110	110
07	113	141	L	133	D	V	D	X	Y	110	123
08	073	157	L	154	D	V	D	X	Y	123	133
09	152	174	L	176	D	V	D	X	N	133	133
10	037	196	L	198	D	V	D	X	N	133	133
11	096	215	L	225	L	V	L	X	Y	133	168
12	227	207	L	213	L	V	L	X	Y	168	191

PARAMETERS

LE = 10
SE = 0
TD = 1
TS = 8
CS = 0
TE = 4
CE = 0
HS = 110
RL = 210



CONDITION: 3 -- STOCHASTIC ERROR

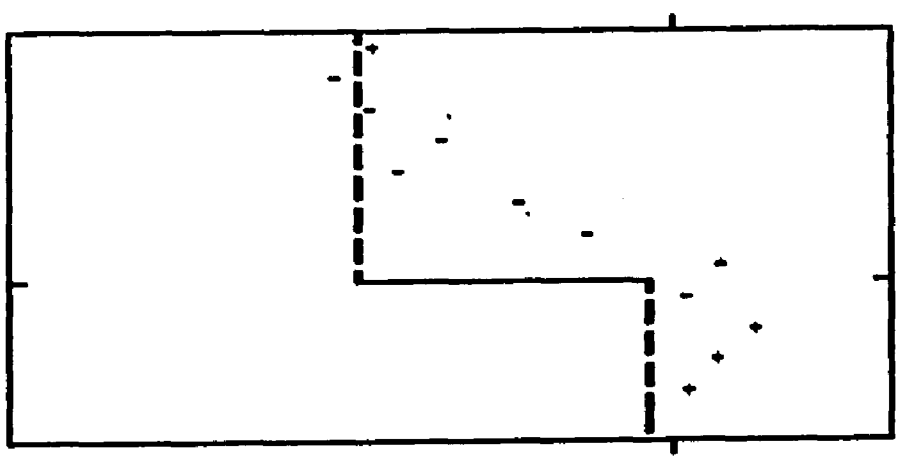
SUBJECT NUM: 1

SUBJECT TAG: 3 PHYSICS EXP 1ST YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	066	D	066	D	V	D	N	N	110	110
01	115	115	L	115	D	B	L	N	N	110	110
02	023	103	D	103	D	V	D	V	N	110	110
03	092	114	L	114	D	V	D	N	N	110	110
04	191	137	L	137	D	V	D	N	N	110	110
05	033	123	D	123	D	V	D	N	N	110	110
06	011	161	L	161	D	V	D	N	N	110	110
07	150	183	L	183	D	V	D	V	N	110	110
08	258	225	L	225	L	B	L	X	Y	110	202
09	210	214	L	214	L	B	D	N	N	202	202
10	156	236	L	236	L	B	L	B	N	202	202
11	052	224	L	224	L	B	L	X	N	202	202
12	100	215	D	215	L	B	L	X	N	202	202

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

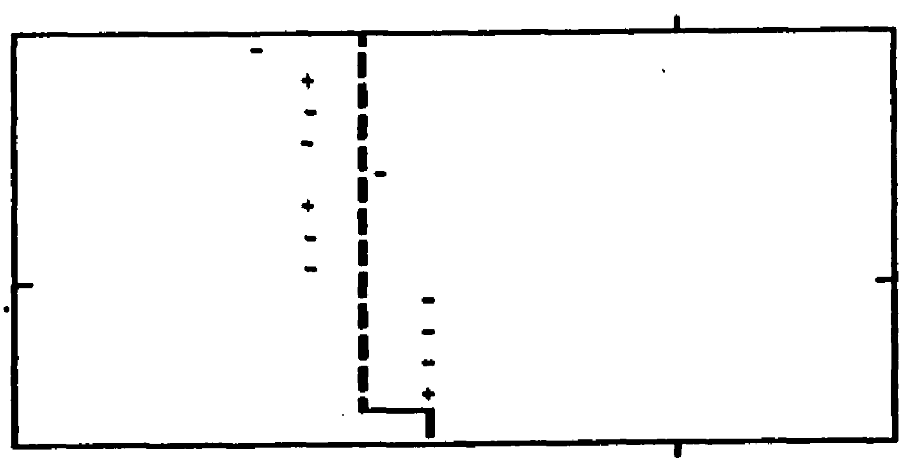
SUBJECT NUM: 2

SUBJECT TAB: 12 PHYSICS EXP 5TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	063	029	D	029	D	B	L	B	Y	110	109
01	145	077	D	077	D	V	D	N	N	110	110
02	211	093	D	093	D	B	L	B	N	110	110
03	224	094	D	094	D	B	D	N	N	110	110
04	087	093	D	093	D	V	D	N	N	110	110
05	116	116	L	116	D	V	D	N	N	110	110
06	127	093	D	093	D	B	L	N	N	110	110
07	201	094	D	094	D	V	D	N	N	110	110
08	245	094	D	094	D	V	D	V	N	110	110
09	121	131	L	131	D	V	D	N	N	110	110
10	101	131	L	131	D	V	D	N	N	110	110
11	185	131	L	131	D	V	D	N	N	110	110
12	110	131	L	131	D	B	L	B	Y	110	131

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

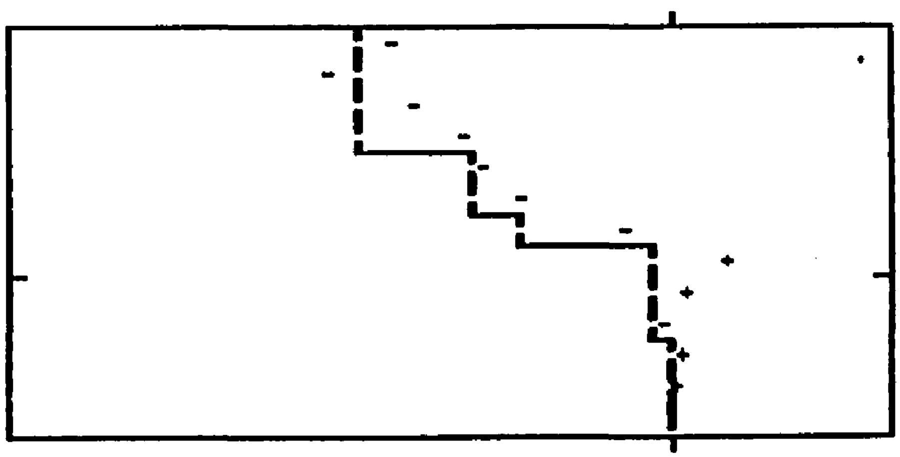
SUBJECT NUM: 3

SUBJECT TAG: 17 PHYSICS EXP 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	132	045	D	045	D	B	L	B	Y	110	109
01	126	121	L	121	D	B	D	B	N	110	110
02	233	101	D	101	D	V	D	N	N	110	110
03	022	128	L	128	D	V	D	N	N	110	110
04	275	144	L	144	D	V	D	N	Y	110	146
05	139	150	L	150	D	V	D	N	N	146	146
06	026	162	L	162	D	V	D	N	Y	146	161
07	082	195	L	195	D	V	D	V	Y	161	203
08	165	227	L	227	L	V	L	X	N	203	203
09	008	214	L	214	L	V	L	N	N	203	203
10	095	207	L	207	D	V	D	V	Y	203	209
11	133	213	L	213	L	V	L	X	N	209	209
12	125	211	L	211	L	B	L	X	N	209	209

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

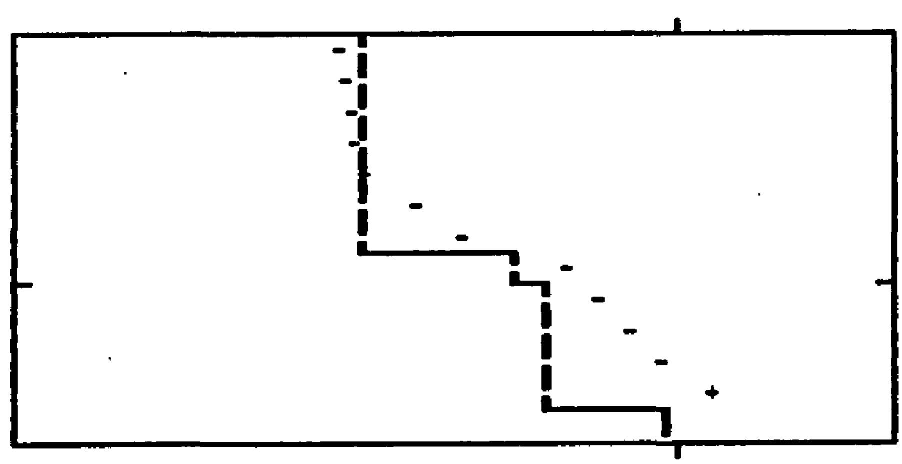
SUBJECT NUM: 4

SUBJECT TAG: 19 PHYSICS EXP 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	031	D	031	D	V	D	V	Y	110	108
01	120	103	D	103	D	V	D	V	N	110	110
02	228	105	D	105	D	V	D	N	N	110	110
03	114	107	D	107	D	V	D	N	N	110	110
04	110	108	D	108	D	V	D	N	N	110	110
05	002	111	D	111	D	B	D	N	N	110	110
06	274	127	L	127	D	B	D	N	N	110	110
07	105	142	L	142	D	V	D	V	Y	110	158
08	024	175	L	175	D	V	D	X	Y	158	168
09	255	185	L	185	D	V	D	N	N	168	168
10	191	195	L	195	D	V	D	N	N	168	168
11	139	205	L	205	D	V	D	N	N	168	168
12	127	221	L	221	L	V	L	V	Y	168	206

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

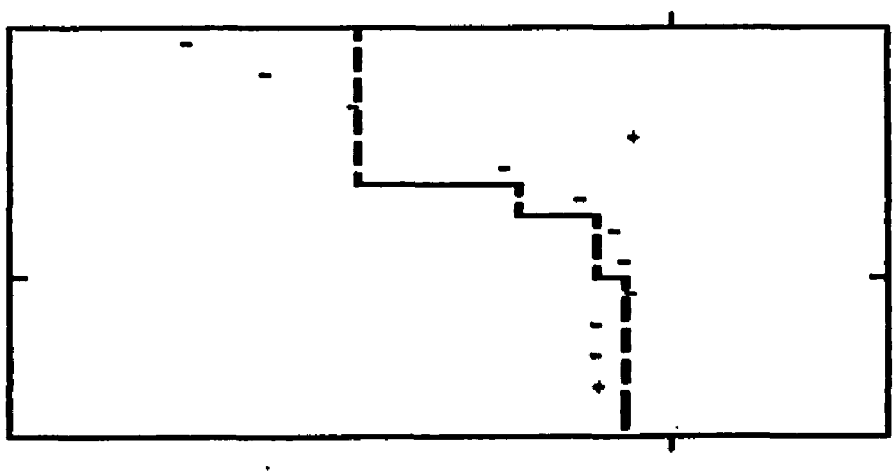
SUBJECT NUM: 5

SUBJECT TAG: 4 CHEMISTRY ORG 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	247	018	D	018	D	V	D	V	Y	110	110
01	043	056	D	056	D	V	D	N	N	110	110
02	105	081	D	081	D	V	D	N	N	110	110
03	109	109	D	109	D	V	D	N	N	110	110
04	122	198	L	198	D	B	L	N	N	110	110
05	118	157	L	157	D	V	D	V	Y	110	161
06	084	181	L	181	D	V	D	V	Y	161	186
07	154	192	L	192	D	V	D	X	N	186	186
08	215	195	L	195	D	V	D	X	Y	186	195
09	005	197	L	197	D	B	D	B	N	195	195
10	276	186	D	186	D	V	D	X	N	195	195
11	255	186	D	186	D	V	D	X	N	195	195
12	267	187	D	187	D	B	L	X	N	195	195

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

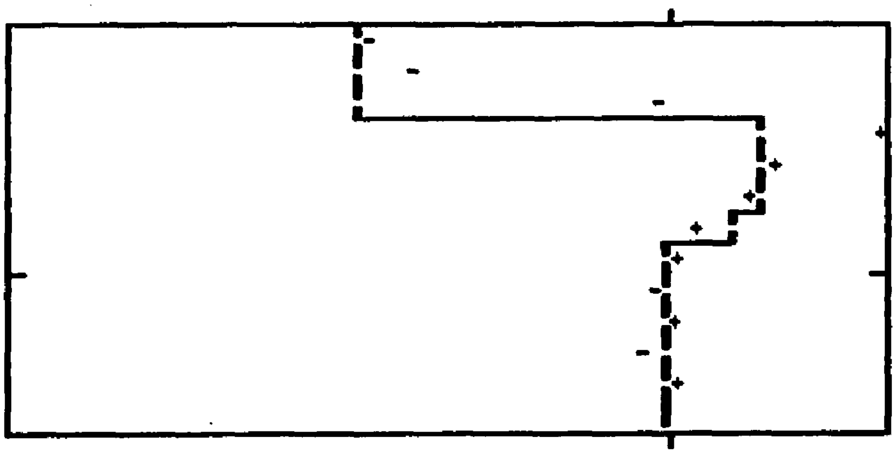
SUBJECT NUM: 6

SUBJECT TAG: 5 CHEMISTRY ANALYT 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	049	D	049	D	V	D	V	Y	110	110
01	120	114	L	114	D	V	D	N	N	110	110
02	047	128	L	128	D	B	D	N	N	110	110
03	092	206	L	206	D	V	D	V	Y	110	238
04	113	277	L	277	L	B	L	N	N	238	238
05	062	243	D	243	L	V	L	N	N	238	238
06	235	235	D	235	L	V	L	N	Y	238	229
07	105	218	D	218	L	V	L	V	Y	229	208
08	024	212	L	212	L	V	L	X	N	208	208
09	255	205	D	205	D	V	D	N	N	208	208
10	191	211	D	211	L	V	L	N	N	208	208
11	139	201	D	201	D	V	D	N	N	208	208
12	127	212	L	212	L	V	L	N	N	208	208

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

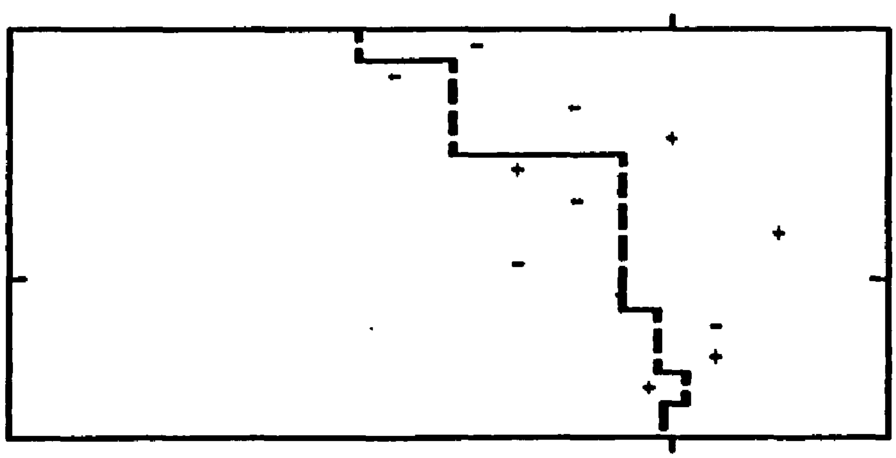
SUBJECT NUM: 7

SUBJECT TAG: 24 CHEMISTRY ORG 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	047	D	047	D	V	D	V	Y	110	105
01	120	148	L	148	D	V	D	N	Y	110	140
02	047	122	D	122	D	B	D	B	N	140	140
03	095	179	L	179	D	V	D	N	N	140	140
04	113	210	D	210	D	B	L	B	Y	140	194
05	222	161	D	161	D	B	L	X	N	194	194
06	183	180	D	180	D	V	D	X	N	194	194
07	196	244	L	244	L	B	L	X	N	194	194
08	032	161	D	161	D	V	D	X	N	194	194
09	092	194	L	194	D	B	D	N	Y	194	205
10	023	224	L	224	L	B	D	B	N	205	205
11	004	224	L	224	L	V	L	X	Y	205	214
12	234	203	L	203	D	V	D	X	Y	214	207

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

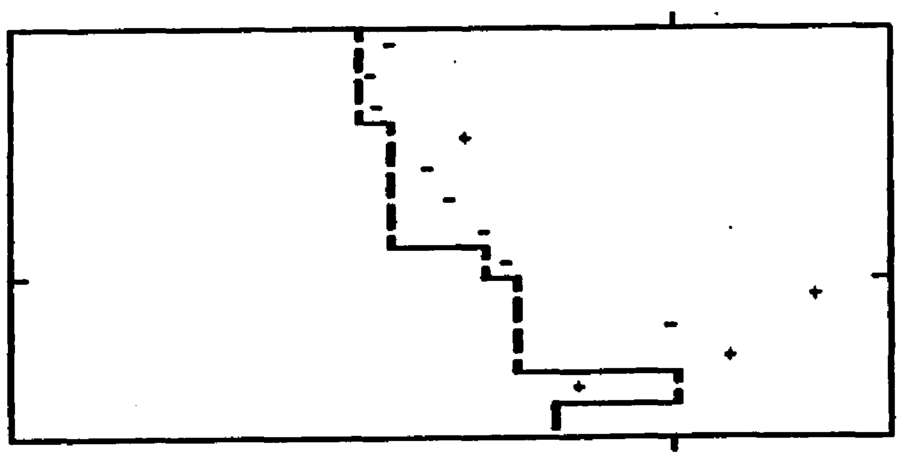
SUBJECT NUM: 8

SUBJECT TAG: 28 BIOLOGY BIOCHEM POST D

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	066	D	066	D	V	D	V	Y	110	107
01	120	120	L	120	D	V	D	N	N	110	110
02	047	114	L	114	D	B	D	B	N	110	110
03	095	116	L	116	D	V	D	N	Y	110	120
04	113	144	L	144	D	B	L	N	N	120	120
05	062	132	L	132	D	V	D	N	N	120	120
06	235	139	L	139	D	V	D	V	N	120	120
07	074	150	D	150	D	V	D	X	Y	120	150
08	024	157	L	157	D	V	D	X	Y	150	160
09	255	255	L	255	L	V	L	N	N	160	160
10	191	209	L	209	D	V	D	N	N	160	160
11	139	228	L	228	L	V	L	V	Y	160	211
12	192	180	D	180	D	B	L	X	Y	211	172

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

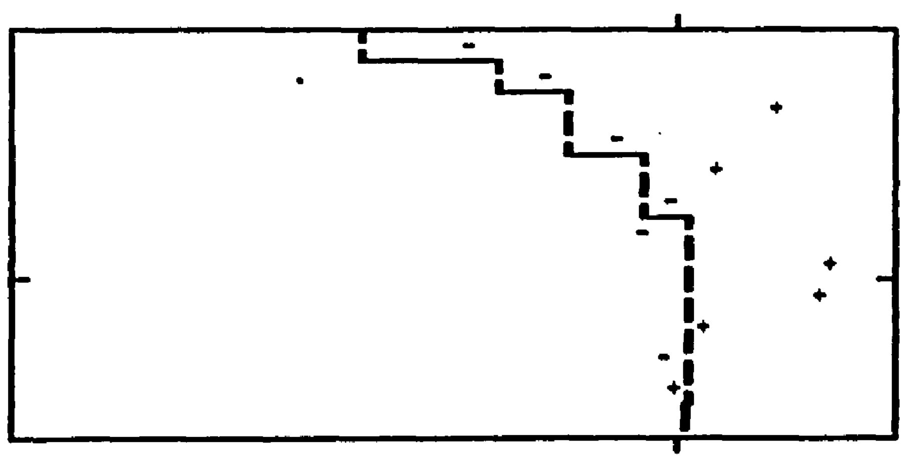
SUBJECT NUM: 9

SUBJECT TAG: 31 BIOLOGY MICRO POST D

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	047	D	047	D	V	D	V	Y	110	110
01	120	144	L	144	D	V	D	N	Y	110	153
02	047	168	L	168	D	B	D	N	Y	153	175
03	092	241	L	241	L	V	L	N	N	175	175
04	191	191	D	191	D	V	D	N	Y	175	199
05	033	222	L	222	L	V	L	N	N	199	199
06	011	208	L	208	D	V	D	N	Y	199	213
07	150	199	D	199	D	V	D	V	N	213	213
08	258	258	L	258	L	B	L	B	N	213	213
09	255	255	L	255	L	V	L	N	N	213	213
10	191	218	L	218	L	V	L	V	N	213	213
11	114	206	D	206	D	V	D	X	N	213	213
12	192	209	D	209	D	B	L	X	Y	213	212

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 3 - STOCHASTIC ERROR

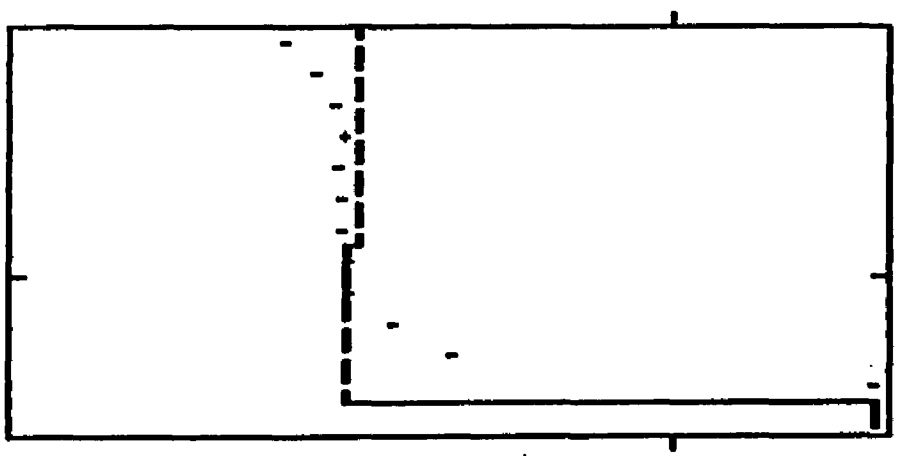
SUBJECT NUM: 10

SUBJECT TAG: 34 BIOLOGY BIOCHEM 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	081	067	D	067	D	V	D	V	Y	110	107
01	120	087	D	087	D	V	D	N	N	110	110
02	047	097	D	097	D	B	D	N	N	110	110
03	092	103	D	103	D	V	D	V	N	110	110
04	113	106	L	106	D	B	L	N	N	110	110
05	062	104	L	104	D	V	D	N	N	110	110
06	235	105	D	105	D	V	D	N	N	110	110
07	105	105	D	105	D	V	D	N	Y	110	106
08	157	107	L	107	D	B	D	B	N	106	106
09	277	107	L	107	D	V	D	N	N	106	106
10	035	121	L	121	D	V	D	V	N	106	106
11	185	140	L	140	D	V	D	X	N	106	106
12	047	274	L	274	L	B	D	X	Y	106	274

PARAMETERS

LE = 0
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

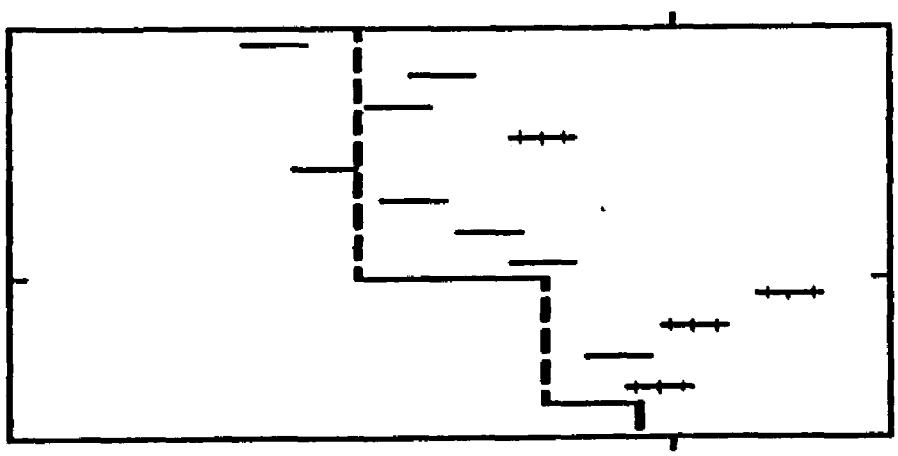
SUBJECT NUM: 1

SUBJECT TAG: 1 PHYSICS EXP 2ND YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	053	D	051	D	V	D	V	Y	110	109
01	121	084	D	092	D	V	D	N	N	110	110
02	054	137	L	127	D	B	D	B	N	110	110
03	098	123	L	130	D	V	D	N	N	110	110
04	115	169	D	178	D	B	L	N	N	110	110
05	067	100	D	103	D	V	D	N	N	110	110
06	228	128	D	118	D	V	D	N	N	110	110
07	107	152	D	159	D	V	D	V	N	110	110
08	032	169	L	159	D	V	D	X	Y	110	169
09	247	247	L	249	L	V	L	N	N	169	169
10	187	217	L	218	L	V	L	N	N	169	169
11	139	193	L	195	D	V	D	V	N	169	169
12	188	206	L	215	L	B	L	X	Y	169	199

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

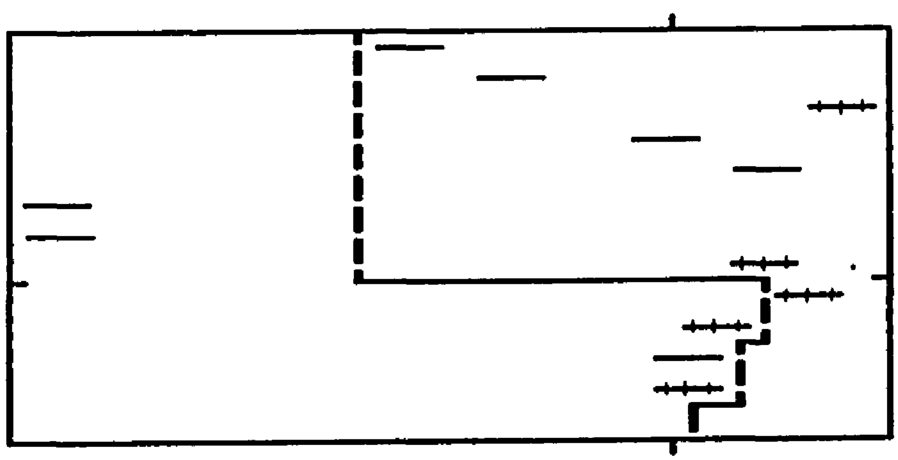
SUBJECT NUM: 2

SUBJECT TAG: 2 PHYSICS EXP 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H' I	H F
D1	085	122	L	120	D	V	D	V	Y	110	110
01	121	127	L	135	D	V	D	V	N	110	110
02	222	159	L	169	D	V	D	N	N	110	110
03	116	264	L	266	L	V	L	N	N	110	110
04	112	208	L	206	D	V	D	N	N	110	110
05	011	240	L	238	L	B	D	N	N	110	110
06	264	015	L	021	D	B	D	B	N	110	110
07	079	016	L	008	D	V	D	X	N	110	110
08	032	239	L	229	L	V	L	X	Y	110	239
09	247	253	L	255	L	V	L	N	N	239	239
10	187	224	D	225	L	V	L	V	Y	239	231
11	116	215	D	205	D	V	D	X	N	231	231
12	188	215	D	224	L	B	L	X	Y	231	216

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

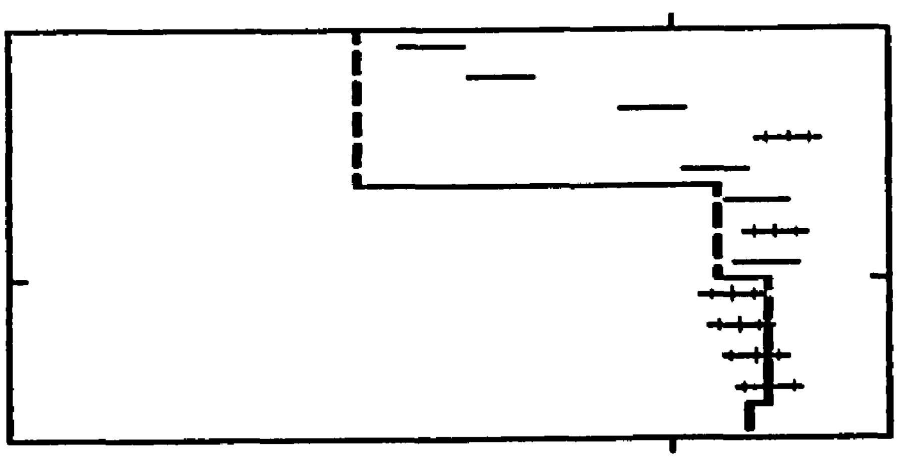
SUBJECT NUM: 3

SUBJECT TAB: 16 PHYSICS EXP 6TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	059	D	057	D	V	D	V	Y	110	108
01	121	134	L	142	D	V	D	V	N	110	110
02	222	156	L	166	D	V	D	N	N	110	110
03	116	204	L	206	D	V	D	N	N	110	110
04	112	247	L	245	L	V	L	N	N	110	110
05	011	224	D	222	L	B	D	N	Y	110	224
06	264	237	L	243	L	B	D	N	N	224	224
07	107	243	L	250	L	V	L	N	N	224	224
08	156	240	L	230	L	B	D	B	Y	224	240
09	267	229	D	228	L	V	L	N	N	240	240
10	043	232	L	232	L	V	L	N	N	240	240
11	062	237	L	235	L	V	L	N	N	240	240
12	055	241	L	238	L	B	L	B	Y	240	234

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

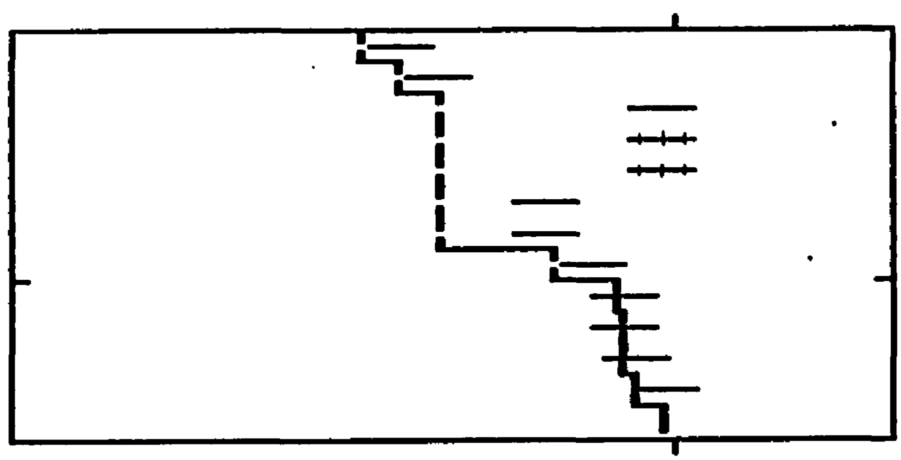
SUBJECT NUM: 4

SUBJECT TAG: 20 PHYSICS EXP 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	018	023	D	026	D	B	D	B	Y	110	109
01	028	123	L	118	D	V	D	V	Y	110	122
02	199	135	L	132	D	V	D	N	Y	122	135
03	163	206	L	200	D	V	D	N	N	135	135
04	144	206	D	215	L	V	L	N	N	135	135
05	085	206	L	214	L	V	L	N	N	135	135
06	212	169	L	176	D	V	D	N	N	135	135
07	019	169	D	176	D	V	D	N	Y	135	171
08	226	184	D	176	D	V	D	V	Y	171	191
09	019	194	D	187	D	B	D	N	Y	191	193
10	163	194	D	195	D	V	D	N	N	193	193
11	110	198	L	191	D	V	D	N	Y	193	197
12	046	207	L	205	D	V	D	V	Y	197	206

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

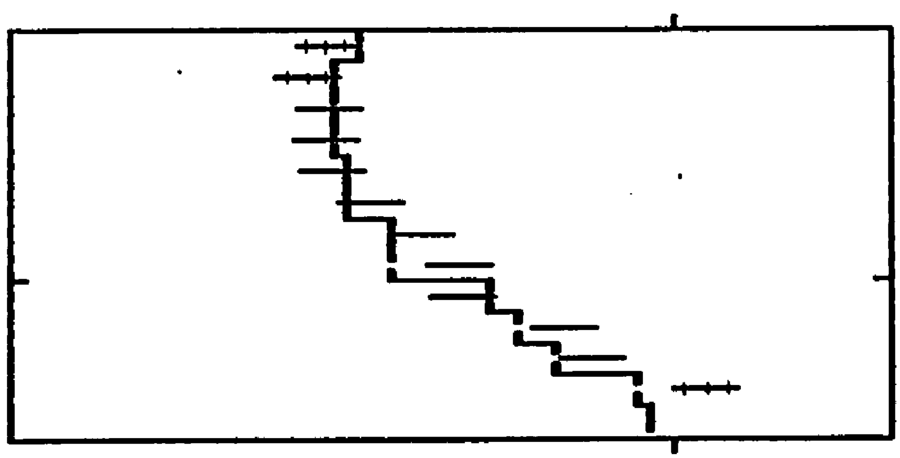
SUBJECT NUM: 5

SUBJECT TAG: 10 CHEMISTRY ORG 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	127	027	D	023	D	V	D	V	Y	110	111
01	101	101	D	109	D	B	L	N	Y	110	102
02	030	094	D	084	D	B	L	B	N	102	102
03	120	101	L	092	D	B	D	N	N	102	102
04	041	100	L	105	D	V	D	N	Y	102	106
05	042	102	D	104	D	B	D	N	N	106	106
06	030	114	L	107	D	V	D	V	Y	106	120
07	100	130	L	123	D	V	D	X	N	120	120
08	148	142	L	151	D	V	D	X	Y	120	151
09	024	143	D	142	D	V	D	N	Y	151	160
10	080	175	L	170	D	V	D	N	Y	160	172
11	163	184	L	175	D	V	D	V	Y	172	198
12	156	220	L	214	L	V	L	X	Y	198	202

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

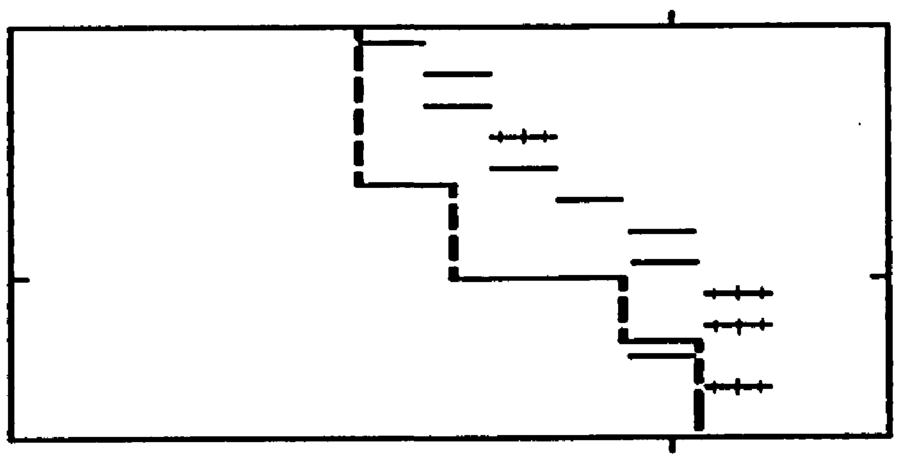
SUBJECT NUM: 6

SUBJECT TAG: 15 CHEMISTRY INORG 3RD YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	144	042	D	043	D	V	D	V	Y	110	112
01	114	121	L	113	D	V	D	V	N	110	110
02	175	142	L	145	D	V	D	N	N	110	110
03	144	142	D	150	D	V	D	N	N	110	110
04	195	163	D	170	D	B	L	B	N	110	110
05	100	163	D	168	D	V	D	X	Y	110	140
06	213	184	D	182	D	V	D	X	N	140	140
07	166	207	D	204	D	V	D	X	N	140	140
08	103	208	D	205	D	V	D	X	Y	140	194
09	229	231	D	225	L	V	L	N	N	194	194
10	212	231	D	233	L	V	L	V	Y	194	218
11	015	207	D	206	D	V	D	X	N	218	218
12	016	231	L	225	L	V	L	X	N	218	218

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

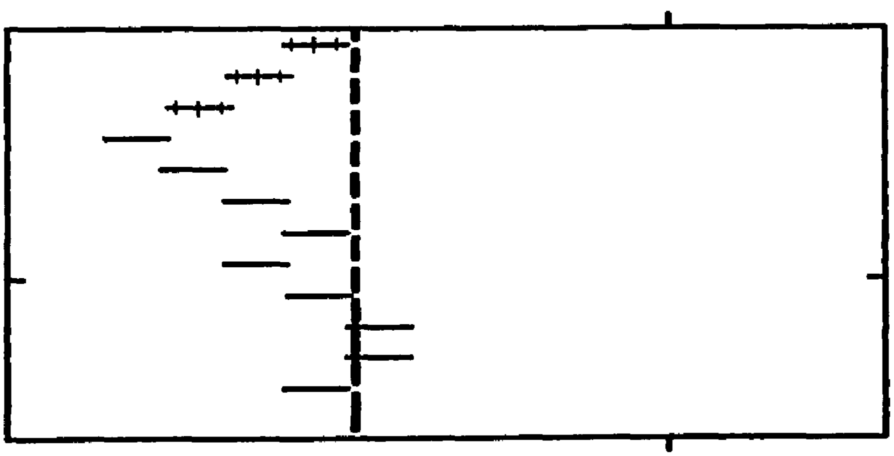
SUBJECT NUM: 7

SUBJECT TAG: 27 CHEMISTRY ANALYT 3RD Y

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	127	052	D	048	D	V	D	V	Y	110	108
01	101	098	D	106	D	B	L	N	N	110	110
02	030	080	D	070	D	B	L	N	N	110	110
03	090	061	D	053	D	B	L	N	N	110	110
04	127	041	D	032	D	B	D	N	N	110	110
05	159	059	D	052	D	V	D	V	N	110	110
06	030	079	D	072	D	V	D	N	N	110	110
07	265	098	D	098	D	V	D	N	N	110	110
08	139	079	D	086	D	V	D	V	N	110	110
09	024	099	D	098	D	V	D	N	N	110	110
10	080	118	D	113	D	V	D	N	N	110	110
11	163	118	D	109	D	V	D	N	N	110	110
12	017	098	D	100	D	V	D	V	N	110	110

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

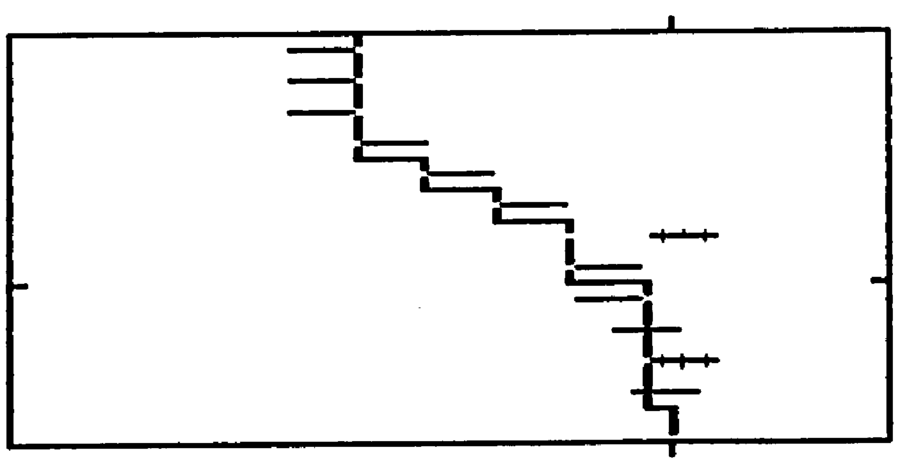
SUBJECT NUM: 8

SUBJECT TAG: 32 BIOLOGY BIOCHEM 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	051	D	049	D	V	D	V	Y	110	107
01	121	099	L	107	D	V	D	N	N	110	110
02	054	099	D	089	D	B	D	N	N	110	110
03	096	099	D	091	D	V	D	N	N	110	110
04	187	122	L	125	D	V	D	V	Y	110	131
05	078	143	L	136	D	V	D	N	Y	131	154
06	264	166	L	172	D	B	D	N	Y	154	177
07	107	214	L	221	L	V	L	N	N	177	177
08	156	190	D	180	D	B	D	B	Y	177	202
09	267	190	L	189	D	V	D	N	N	202	202
10	043	202	D	202	D	V	D	N	N	202	202
11	062	214	L	212	L	V	L	V	N	202	202
12	054	208	L	201	D	B	D	X	Y	202	210

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 .- MEASUREMENT AND STOCHASTIC ERROR

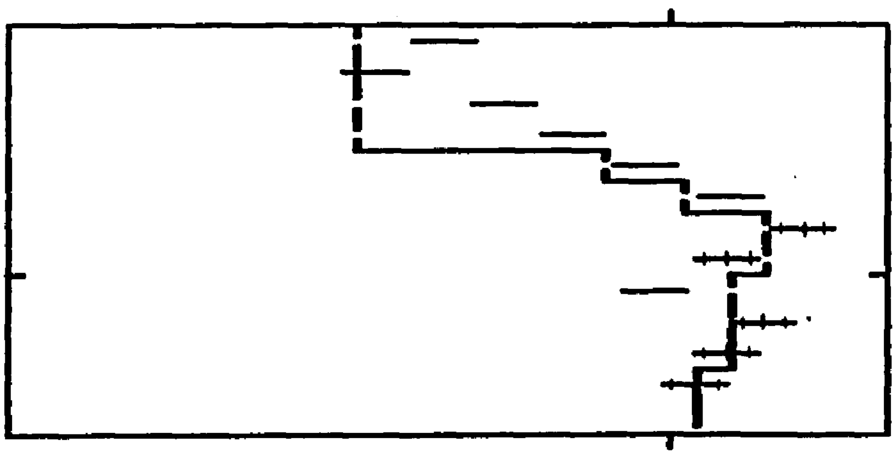
SUBJECT NUM: 9

SUBJECT TAG: 38 BIOLOGY MICRO 4TH YR

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	067	D	065	D	V	D	V	Y	110	105
01	121	138	L	146	D	V	D	N	N	110	110
02	054	116	D	106	D	B	D	N	N	110	110
03	096	157	L	149	D	V	D	N	N	110	110
04	187	179	L	182	D	V	D	V	Y	110	189
05	078	202	L	195	D	V	D	N	Y	189	214
06	264	229	L	235	L	B	D	N	Y	214	240
07	107	252	L	259	L	V	L	V	N	240	240
08	032	228	D	218	L	V	L	X	Y	240	229
09	247	205	D	207	D	V	D	V	N	229	229
10	127	239	L	235	L	V	L	X	N	229	229
11	116	228	L	218	L	V	L	X	Y	229	218
12	188	218	L	227	L	B	L	X	N	218	218

PARAMETERS

LE = 10
BE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



CONDITION: 4 - MEASUREMENT AND STOCHASTIC ERROR

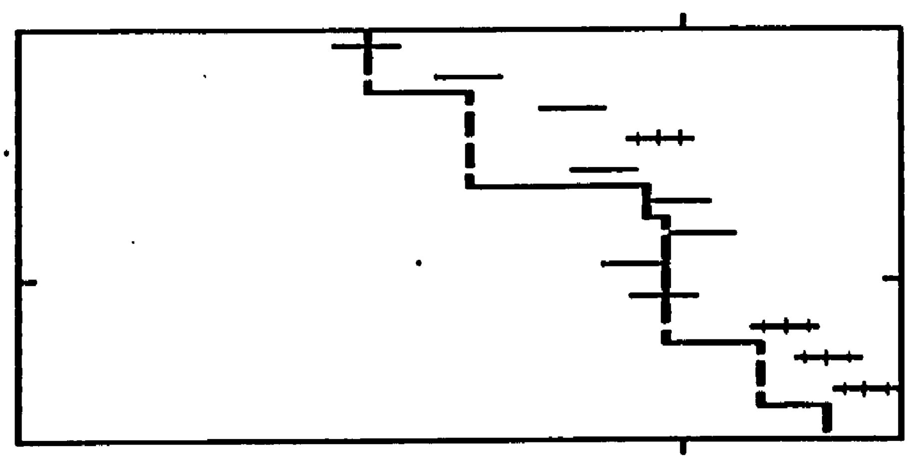
SUBJECT NUM: 10

SUBJECT TAG: 40 BIOLOGY MICRO POST DOC

T N	T I	T F	T P	T L	T V	P I	T A	P C	H M	H I	H F
D1	085	057	D	055	D	V	D	V	Y	110	108
01	121	110	L	118	D	V	D	N	N	110	110
02	054	142	L	132	D	B	D	B	Y	110	142
03	098	175	L	182	D	V	D	N	N	142	142
04	115	203	L	212	L	B	L	N	N	142	142
05	067	185	D	188	D	V	D	N	Y	142	198
06	228	208	L	198	D	V	D	V	Y	198	204
07	079	216	L	208	D	V	D	X	N	204	204
08	032	195	L	185	D	V	D	X	N	204	204
09	247	204	L	206	D	V	D	N	N	204	204
10	187	242	L	243	L	V	L	V	Y	204	234
11	116	256	L	246	L	V	L	X	N	234	234
12	188	268	L	277	L	B	L	X	Y	234	255

PARAMETERS

LE = 10
SE = .25
TD = 1
TS = 8
CS = 2
TE = 4
CE = 1
HS = 110
RL = 210



REFERENCES

Anderson, N.H., & Norman, A. Order effects in impression formation in four classes of stimuli. Journal of Abnormal and Social Psychology, 1964, 69, 467-471.

Asch, S. Forming impressions of personality. Journal of Abnormal and Social Psychology, 1946, 41, 258-290.

Bakan, D. On method. San Francisco: Jossey-Bass, 1967.

Bartlett, F.C. Remembering: A study in experimental and social psychology. Cambridge: Cambridge University Press, 1932.

Bransford, J.D., & Franks, J.J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Chapman, L.J. Illusory correlation in observational report. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 151-155.

Chapman, L.J., & Chapman, J.P. Genesis of popular but erroneous psychodiagnostic observations. Journal of Abnormal Psychology, 1967, 72, 193-204.

Chapman, L.J., & Chapman, J.P. Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology, 1969, 74, 271-280.

Cofer, C.N. Constructive processes in memory. American Scientist, 1973, 61, 537-543.

Copi, I.M. Introduction to logic (5th Ed.). New York: Macmillan, 1978.

D'Andrade, R.G. Trait psychology and componential analysis. American Anthropologist, 1965, 67 (5), 215-228.

D'Andrade, R.G. Memory and the assessment of behavior. In H. Blalock (Ed.), Measurement in the social sciences. Chicago: Aldine-Atherton, 1974.

Dailey, C.A. The effects of premature conclusions upon the acquisition of understanding of a person. Journal of Psychology, 1952, 33, 133-152.

Doherty, M.E., Mynatt, C.R., Tweney, R.D., & Schiavo, M.D. Pseudodiagnosticity. Acta Psychologica, 1979, 43, 111-121.

Einhorn, H.J., & Hogarth, R.M. Confidence in judgments: Persistence of the illusion of validity. Psychological Review, 1978, 85, 395-416.

Evans, J. St. B.T. Interpretation and matching bias in a reasoning task. Quarterly Journal of Experimental Psychology, 1972, 24, 193-199.

Evans, J. St. B.T., & Lynch, J.S. Matching bias in the selection task. British Journal of Psychology, 1973, 64, 391-397.

Faraday, M. Faraday's diary: Being the various philosophical notes of experimental investigation made by Michael Faraday during the years 1820-1962... London: G. Bell, 1932-1936 (7 vols. plus index. Ed. by T. Martin).

Fischoff, B. Attribution theory and judgment under uncertainty. In J. Harvey, W. Ickes, and R. Kidd (Eds.), New directions in attribution research. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976, 419-450.

Fischoff, B., & Beyth, R. "I knew it would happen"--remembered probabilities of once-future things. Organizational Behavior and Human Performance, 1975, 13, 1-16.

Fischoff, B., Slovic, P., & Lichtenstein, S. Knowing with certainty: The appropriateness of extreme confidence. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 552-564.

Fisher, S., & Fisher, R. Test of certain assumptions regarding figure drawing analysis. Journal of Abnormal and Social Psychology, 1950, 45, 727-732.

Franks, J.J., & Bransford, J.D. Abstraction of visual patterns. Journal of Experimental Psychology, 1971, 90, 65-74.

Goldberg, L.R. Simple models or simple processes? Some research on clinical judgments. American Psychologist, 1968, 23, 483-496.

Goldberg, L.R., & Rorer, L.G. Learning clinical inference: The results of intensive training on clinicians' ability to diagnose psychosis versus neurosis from the MMPI. Paper presented at the meeting of the Western Psychological Association, Honolulu, June 1965.

Golding, S.L., & Rorer, L.G. Illusory correlation and selective judgment. Journal of Abnormal Psychology, 1972, 80, 249-260.

Goodwin, R.Q., & Wason, P.C. Degrees of insight. British Journal of Psychology, 1972, 63, 205-212.

Greenwald, A.G. On the conceptual disconfirmation of theories. Personality and Social Psychology Bulletin, 1981, 7, 131-137.

Hadamard, J.S. An essay on the psychology of invention in the mathematical field. Princeton: Princeton University Press, 1945.

Hamilton, D.L., & Gifford, R.K. Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. Journal of Experimental Social Psychology, 1976, 12, 392-407.

Hastorf, A.H., & Cantril, H. They saw a game: A case study. Journal of Abnormal and Social Psychology, 1954, 49, 129-134.

Holtzberg, J.D., & Wexler, M. The validity of human form drawings as a measure of personality deviation. Journal of Projective Techniques, 1950, 14, 343-361.

Hempel, C.G. Philosophy of natural science. New Jersey: Prentice-Hall, 1966.

Hume, D. An inquiry concerning human understanding. New York: Bobbs-Merrill, 1955.

Jackson, D.N. An analysis of the generality of trait inferential networks. (Research Bulletin No. 78) London: University of Western ONtario, Department of Psychology, April 1968.

Johnson-Laird, P.N., Legrenzi, P., & Sonino Legrenzi, M. Reasoning and a sense of reality. British Journal of Psychology, 1972, 63, 395-400.

Johnson-Laird, P.N., & Wason, P.C. A theoretical analysis of insight into a reasoning task. Cognitive Psychology, 1970, 1, 134-138.

Kahneman , D., & Tversky, A. Subjective probability: A judgment of representativeness. Cognitive Psychology, 1972, 3, 430-454.

Kahneman , D., & Tversky, A. On the psychology of prediction. Psychological Review, 1973, 80, 237-251.

Kanouse, D.E. Language, labeling, and attribution. In E.E. Jones, D.E. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins, and B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. Morristown, New Jersey: General Learning Press, 1972.

Kern, L.H. Confirmation bias: An integration of the literature. Unpublished manuscript, The Ohio State University, 1981.

Kern, L.H., & Doherty, M.E. "Pseudodiagnosticity" in an idealized medical problem-solving environment. Journal of Medical Education, 1982, 57, 100-104.

Kern, L.H., Mirels, H.L., & Hinshaw, V.G. Scientists' understanding of propositional logic: An experimental investigation. Social Studies of Science, 1982, in press.

Koltuv, B.B. Some characteristics of intrajudge trait intercorrelations. Psychological Monographs, 1962, 76, (33, Whole No. 552).

Loftus, E.F. Leading questions and the eyewitness report. Cognitive Psychology, 1975, 7, 560-572.

Loftus, E.F., & Palmer, J. Reconstruction of automobile destruction. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 585-589.

Mahoney, M.J. Scientist as subject: The psychological imperative. Cambridge, Mass.: Ballinger, 1976.

Mahoney, M.J. Publication prejudices: An experimental study of confirmatory bias in the peer review system. Cognitive Therapy and Research, 1977, 1(2), 161-175.

Mahoney, M.J., & DeMonbreun, B.G. Psychology of the scientist: An analysis of problem-solving bias. Cognitive Therapy and Research, 1977, 1(3), 229-238.

Mahoney, M.J., & Kimper, T.P. From ethics to logic: A survey of scientists. In Mahoney, M.J. Scientist as subject: The psychological imperative. Cambridge, Mass.: Ballinger, 1976.

Markowitz, L.M., & Tweney, R.D. An investigation of the behavior of mathematicians engaged in testing a conjecture. Unpublished manuscript. Bowling Green State University, 1981.

Meehl, P.E. Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press, 1954.

Meehl, P.E. The cognitive activity of the clinician. American Psychologist, 1960, 15, 19-27.

Meehl, P.E. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. Journal of Consulting and Clinical Psychology, 1978, 46(4), 806-834.

Mirels, H.L. Implicit personality theory and inferential illusions. Journal of Personality, 1976, 44, 467-487.

Mischel, W. Personality and assessment. New York: Wiley, 1968.

Mischel, W. Towards a cognitive social learning reconceptualization of personality. Psychological Review, 1973, 80, 252-283.

Mischel, W. On the future of personality measurement. American Psychologist, 1977, 32, 246-254.

Mischel, W., Ebbesen, & Zeiss. Determinants of selective memory about the self. Journal of Consulting and Clinical Psychology, 1976, 44, 92-103.

Mitroff, I.I. The subjective side of science: A philosophical inquiry into the psychology of the Apollo moon scientists. New York: Elsevier, 1974.

Mynatt, C.R., Doherty, M.E., & Tweney, R.D. Confirmation bias in a simulated research environment: An experimental study of scientific inference. Quarterly Journal of Experimental Psychology, 1977, 29, 85-95.

Mynatt, C.R., Doherty, M.E., & Tweney, R.D. Consequences of confirmation and disconfirmation in a simulated research environment. Quarterly Journal of Experimental Psychology, 1978, 30, 395-406. (a)

Mynatt, C.R., Doherty, M.E., & Tweney, R.D. On the psychology of science. Paper presented at the First Leonard Memorial Conference on Scientific Discovery, Reno, Nevada, 1978. (b)

Newcomb, T. An experiment designed to test the validity of a rating technique. Journal of Educational Psychology, 1931, 22, 279-289.

Newman, R. Unpublished data. Department of Physics, University of California at Irvine, 1982.

Nisbett, R.E., & Lemley, R.N. The evil that men do can be diluted, the good cannot. Unpublished manuscript, University of Michigan, 1979.

Nisbett, R.E., & Ross, L. Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs, N.J.: Prentice Hall, 1980.

Nisbett, R.E., & Zukier, H. The dilution effect: Producing "regressive" predictions by exposure to nondiagnostic information. Unpublished manuscript, University of Michigan, 1979.

Norman, W.T., & Goldberg, L.R. Raters, ratees, and randomness in personality structure. Journal of Personality and Social Psychology, 1966, 4, 681-691.

Oskamp, S. Overconfidence in case-study judgments. Journal of Consulting Psychology, 1965, 29, 261-265.

Passini, F.T., & Norman, W.T. A universal conception of personality structure. Journal of Personality and Social Psychology, 1966, 4, 44-49.

Peterson, C.R., & Ducharme, W.M. A primacy effect in subjective probability revision. Journal of Experimental Psychology, 1967, 73(1), 61-65.

Phillips, L.D., & Edwards, W. Conservatism in a simple probability inference task. Journal of Experimental Psychology, 1966, 72, 346-354.

Pitz, G.F., Downing, L., & Reinhold, H. Sequential effects in the revision of subjective probabilities. Canadian Journal of Psychology, 1967, 21, 381-393.

Platt, J.R. Strong inference. Science, 1964, 146, 347-353.

Polya, G. Mathematics and plausible reasoning Vol. II: Patterns of plausible inference. Princeton, N.J.: Princeton University Press, 1954.

Popper, K.R. The logic of scientific discovery. London: Hutchinson, 1959.

Popper, K.R. Conjectures and refutations. New York: Basic Books, 1962.

Pruitt, D.G. Informational requirements in making decisions. American Journal of Psychology, 1961, 74, 433-439.

Rosenhan, D.L. On being sane in insane places. Science, 1973, 179, 250-258.

Ross, L., Lepper, M., & Hubbard, M. Perseverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. Journal of Personality and Social Psychology, 1975, 32, 880-892.

Ross, L., Lepper, M.R., Strack, F., & Steinmetz, J. Social explanations and social expectation: Effects of real and hypothetical explanations on subjective likelihood. Journal of Personality and Social Psychology, 1977, 32, 817-829.

Ross, M., & Sicoly, F. Egocentric biases in availability and attribution. Journal of Personality and Social Psychology, 1979, 37, 322-336.

Sarbin, T.R., Taft, R., & Bailey, D.E. Clinical inference and cognitive theory. New York: Holt, Rinehart and Winston, 1960.

Sawyer, J. Measurement and prediction: Clinical and statistical. Psychological Bulletin, 1966, 66, 178-200.

Shweder, R.A. Likeness and likelihood in everyday thought: Magical thinking and everyday judgments about personality. Current Anthropology, 1977, 18(4), 637-658.

Shanteau, J. An additive model for sequential decision making. Journal of Experimental Psychology, 1970, 85, 181-191.

Shanteau, J. Descriptive versus normative models of sequential inference judgment. Journal of Experimental Psychology, 1972, 93, 63-68.

Shanteau, J. Averaging versus multiplying combination rules of inference judgment. Acta Psychologica, 1975, 39, 83-89.

Sines, L.K. The relative contribution of four kinds of data to accuracy in personality assessment. Journal of Consulting Psychology, 1959, 23, 483-492.

Smedslund, J. The concept of correlation in adults. Scandinavian Journal of Psychology, 1963, 4, 165-173.

Snyder, M., Tanke, E.D., & Berscheid, E. Social perception and interpersonal behavior: On the self fulfilling nature of social stereotypes. Journal of Personality and Social Psychology, 1977, 35, 656-666.

Snyder, M., & Uranowitz, S.W. Reconstructing the past: Some cognitive consequences of person perception. Journal of Personality and Social Psychology, 1978, 36, 941-950.

Swensen, C.H. Empirical evaluations of human figure drawings. Psychological Bulletin, 1957, 54, 431-466.

Troutman, C.M., & Shanteau, J. Inferences based on nondiagnostic information. Organizational Behavior and Human Performance, 1977, 19, 43-55.

Tversky, A., & Kahneman, D. The belief in the law of small numbers. Psychological Bulletin, 1971, 76, 105-110.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 1973, 5, 207-232.

Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

Tweney, R. Confirmatory and disconfirmatory heuristics in Michael Faraday's scientific research. Paper presented at the Twenty-second Annual Meeting of the Psychonomic Society, Philadelphia, Pennsylvania, 1981.

Tzeng, O. J-L. An empirical study of the differences between components of a dual-memory system. Unpublished doctoral dissertation, The Pennsylvania State University, 1972.

Ward, W.D., & Jenkens, H.M. The display of information and the judgment of contingency. Canadian Journal of Psychology, 1965, 19, 231-241.

Wason, P.C. On the failure to eliminate hypotheses in a conceptual task. Quarterly Journal of Experimental Psychology, 1960, 12, 129-140.

Wason, P.C. The effect of self-contradiction on fallacious reasoning. Quarterly Journal of Experimental Psychology, 1964, 16, 30-34.

Wason, P.C. Reasoning about a rule. Quarterly Journal of Experimental Psychology, 1968, 20, 273-278.

Wason, P.C., & J. St. B.T. Evans. Dual processes in reasoning? Cognition, 1975, 3, 141-154.

Wason, P.C., & Golding, E. The language of inconsistency. British Journal of Psychology, 1974, 65, 537-546.

Wiggins, J.S. Personality and prediction: Principles of personality assessment. Reading, Mass.: Addison-Wesley, 1973.

Youngblood, J., & Himmelfarb, S. The effects of prior neutral messages on resistance to evaluative communication. Psychonomic Science, 1972, 29, 348-350.

Zangwill, O.L. Remembering revisited. Quarterly Journal of Experimental Psychology, 1972, 24, 123-138.

Zukier, H. On the psychology of the red herrings: The role of non-diagnostic information in prediction. Unpublished manuscript, New School for Social Research, 1979.