1966

Revision of Subjective Probabilities Under a Bayesian Model

Gary A. Sterner
Central Washington University

REVISION OF SUBJECTIVE PROBABILITIES UNDER A BAYESIAN MODEL

A Thesis

Presented to

the Graduate Faculty

Central Washington State College

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Psychology

---

by

Gary A. Sterner

August 1966

APPROVED FOR THE GRADUATE FACULTY

Jack J. Crawford, COMMITTEE CHAIRMAN

Richard B. Morris

Eldon E. Jacobsen

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

THE PROBLEM

The purpose of this study was to examine the degree to which subjective probability judgments conform to the Bayesian model of mathematical probability theory; more specifically, the degree to which subjective probability estimates for intersections of events approximate the product of the subjective judgments for component events. Events were defined as concepts based on the uncertain relationships between experimental cues and classes. In general, the study followed the model of the probabilistic concept formation studies.

The bulk of the literature related to probabilistic concept formation has its origins in the theoretical work of Egon Brunswik (1943,1947,1952,1956). The Probabilistic Functionalism Theory of Brunswik, as well as the Transactional Functionalism Theory of the Ames group, views subjects as operating in a complex environment in which cues associated with objects are not perfectly correlated. Most of our judgments as to whether objects are near or far, for example, rely upon combinations of partially valid cues. Allport (1955), in discussing Brunswik's theory, states that environmental ambiguity adds to the uncertain cue-to-object relationships so that judgments are never better than a probable achievement. An unconscious weighting process is suggested in which the object most frequently associated with the particular cues now presented to the subjects will be the most probable on the present occasion. Brunswik (1943) theorizes that the subject effects a compromise between the cues available on the basis of their respective trustworthiness and makes a "best wager."

The two basic approaches to probabilistic concept formation have involved either correlation and multiple regression or the use of probability theory. The typical correlational approach has been to establish several cue dimensions which have predetermined correlations with a criterion variable. Subjects are presented with cues from the various cue dimensions and their task is to predict the criterion variable. Subjects are provided with feedback and experience with the cue-criterion universe. It is hypothesized that they will learn the correlations between the cues and the criterion variables. The extent to which the cue-criterion relationship is learned is determined by computing the correlations between the subject's responses and each cue dimension.

The paradigm for experiments using probability have generally consisted of presenting subjects with different cues, each of which have a predetermined conditional probability relationship with two or more classes. The concepts to be learned consist of the conditional probabilities of the various classes given each cue. The degree to which these probabilistic rules are learned is then examined on later test trials by having subjects perform tasks utilizing their subjective notions.

Correlational relationships permit an examination of the use of cues for inferences about a group of objects or for repeated inferences about a single object. Whereas, probabilistic relationships allow for the examination of each object as a unique event.

A correlational study by Summers (1962), using color, area, and position of a pointer as cues, and line lengths as a criterion, showed that subjects depended upon the cues in the same rank order as the cues objective validities, but that correlations between the cues and subjects' responses were low.

In another study in which the cues and criteria were considerably more clear, Uhl (1963) obtained impressive results. Three rows of nine colored lights were used as cue dimensions and a row of amber lights in front of the subjects served as the criterion feedback. The correlations between the subjects' responses and the criterion feedback varied as a function of the validity of the cues. The mean correlations were from about .80 when the three cues had equal .57 relationships to the criterion, through .90 when the cue-criterion relationships were .00, .45, and .88, to .99 when they were .00, .01, and .99. The mean correlations between subjects' responses and the cues closely approximated the objective cue-to-criterion correlations.

In general, the results of the correlational studies indicate that subjects can learn to utilize equivocal cues in order to predict related criteria.

One of the earliest experiments using probabilistic cues was performed by Robert Goodow (1954). Airplane silhouettes were classified into one of two classes (X or not-X) on the basis of three cues: wing shape, air scoop location, and tail shape. Subjects were given one cue with a probability of 1.00 on each trial and two cues which indicated the correct answer .67 of the time. The cues in this case were presented singly. On the fifth block of 100 trials subjects' responses followed the 1.00 cue about .89 of the time and the .67 cues about .86 of the time. When subjects were presented with two cues, one a 1.00 cue and the other a .67 cue and both cues pointed to the same answer, the answers were almost always the one indicated, i.e., the results were comparable to the use of a 1.00 cue alone. However, when the 1.00 and .67 cues were in conflict, subjects tended to follow the 1.00 cue only .87 of the time. The conflicting cues seemed to reduce certainty, but it was noted that subjects tended to rely on the 1.00 cue more and more as the number of trials increased.

When subjects were given two cues which were both .67 cues and which both pointed to the same answer, they gave that answer .90 of the time. When these two cues were in conflict, subjects gave the answer indicated by each .50 of the time. These results appear to confirm the notion that subjects can perform reasonably well in a probabilistic concept formation situation.

In a more complex experiment, Beach (1964) had subjects learn to predict to which of three classes each of a set of 120 cards belonged. The predictions were based on the cards' three cue values which were from three nine-valued cue dimensions. After the subjects had been presented a card and had made a judgment about its class membership, they were told the correct class. It was hypothesized that subjects would learn the probability relationships between the multiple cues and the classes. Special cue combinations for which the frequency of association values led to a class choice which was not the one designated as correct by the feedback were scattered throughout the deck of cards. It was found that subjects made persistent errors on these infirming cards during the first part of the experiment. The errors, for the most part, constituted the substitution of the probability based class choice, which indicates that probabilities were being utilized. From the beginning, the subject's responses for the cards which possessed cues predicting the actual class of assignment were increasingly in the direction of the correct class. During the later part of the experiment, evidence of subject recognition occurred and all of the responses tended to be the class given as correct. There were 1440 training trials in the experiment which may have been too many in light of possible card memorization. The number of times the subjects saw a single card was not reported.

The influence of probabilism was also checked after the training trials when subjects were shown test cards possessing combinations of two and three cues. The combinations had not been used during the training period. The subjects were asked to choose a "best bet" for the correct class. The subjects' class choices were compared to the probability distributions computed by adding the percentage of times the cues were associated with each class during the training trials. For 10 test cards out of 14, subjects chose the class having the highest past association, i.e., the "best bet." A second experiment using 12 valued cue dimensions and five classes yielded similar results. These findings support the notion that subjects can learn probability relationships.

While the results of the probabilistic concept formation studies indicate that subjects can become quite accurate over time, they do not expose the process of learning the relationships, i.e., the mechanics of the compromise. When the subject is first gaining knowledge about the experimental universe, his judgments cannot be expected to conform to the established experimental probability relationships. If these inaccurate subjective estimates reflect the subject's actual subjective evaluation of the cue-class relationships, and if subjects are good intuitive statisticians, as Brunswik has suggested, then we might expect to find that the inter-relations among subjective probability estimates would be consistent in terms of some model. That is, we would expect an orderly compromise on the basis of a subject's subjective estimates of the data. The specific manner in which we could expect to find these consistencies among subjective probabilities is open to question.

A new and rapidly developing area of research and application viewing men as probabilistic information processors is based on Bayesian statistics. Bayes's theorem is said to be the appropriate normative model defining the manner in which humans function as intuitive statisticians. A normative model, according to Edwards, Lindman, and Phillips (1965) is "a set of rules specifying what people should do; a normative model for decision making then, specifies what decision you should make." The equation is:

$$P(H|D) = \frac{P(D|H)\ P(H)}{P(D)} \tag{1}$$

Where the P(H) is the prior probability of the hypothesis under consideration being correct; the P(D| H) is the probability of the data being associated with the hypotheses, the P(D) is the prior probability of the data occuring, and the P(H|D) is the revised posterior probability that the hypothesis is correct, given the data associated with it.

Bayes's theorem, which permits the revision of opinions on the basis of new information, has been used very little in the past. This is largely because of the difficulty in making probability judgments prior to the collection of data.

According to Stilson (1966), the Bayesian approach has been reappraised within the past decade and a group of statisticians and psychologists referred to as the neo-Bayesians have emerged. The neo-Bayesians obtain the prior probabilities by looking to the knowledge, experience, and intuition of an individual to assign subjective prior probabilities. Stilson (1966) raises the question about the mathematical rules for subjective probabilities, i.e., when subjective and objective probabilities are combined in one mathematical formula, as the neo-Bayesian does when he applies Bayes Theorem, the assumption is made that subjective probabilities combine according to the same rules of mathematical probability theory as the objective probabilities do. This same question can be raised in the event that subjective probabilities alone are combined in the manner prescribed by Bayes's Theorem. By examining the degree to which subjective probability judgments conform to the Bayesian model, the present study assumes that subjective probabilities obey the mathematic rules. Studies investigating this assumption submit subjective probabilities to the specific rules of probability theory. Two such rules are: (1) the sum of all outcome probabilities in a given sample space must be unity; (2) the sum of the probabilities of the outcomes contained in an event equals the probability of that event.

Subjective probability, as used here, refers to what Edwards, <u>et al</u> (1965), call the "personalistic view." An individual's statement of probability is said to be a matter of opinion and is associated with that person only. This definition of subjective probability points to the specific person only insofar as his experience or data differs from the next person. It is implied that men have in common a method of information processing so that inferences from completely adequate data may be considered to be objective. While a person's initial opinion may differ from his neighbor's, Edward <u>et al</u> (1965) suggests that both opinions may be transformed to a series of relevant observations so as to become nearly indistinguishable. Edwards et al (1965) states that "the personalistic approach permits just as meaningful a discussion of the probability of a unique event as of the probability of a repeatable event." The present study and the related literature deal with the repeatable event in that there is a relative frequency link between data and hypothesis. Other studies involving subjective probability, according to Edwards (1966), examine plausibility links, based on general verbal knowledge about the unique event.

Beach (1966) reports a study utilizing Bayes's theorem as the appropriate normative model for how subjects revise opinions under the paradigm of the probabilistic concept formation studies. In experiment (1) by Beach (1966), revision consistency was examined. Revision consistency, as used by Beach, refers to the degree of consistency between subjects' paired cue estimates, and the Bayesian revision of their single cue estimates. More specifically, Beach defines revision consistency as the degree to which the two sides of Bayes's theorem as shown in equation (1) approach equality when subjects' prior subjective probabilities are revised to obtain posterior subjective probabilities.

The equation that Beach used to revise single cue estimates involves a form of Bayes's Theorem which is based on the experimental procedure as well as the statistical structure of the experimental universe. The subjects in the experiment were made aware of the fact that the classes were equally represented and therefore equally probable. The known equal probability of classes plus the fact that the probability of a cue is, by definition, constant across classes, enabled Beach to define the Bayesian relationship between a single cue estimate of probability and the probability of the single cue being associated with the class as:

$$P\ (H|d) \propto P\ (d|H) \qquad (2)$$

which states that the subjective probability of a class given a cue is proportional to the subjective probability of the cue given the class. When there are several cues available and they are independent, the Bayesian equation for the multiple cue relationship is:

$$P\ (H|d_1, d_2, d_3) \propto P\ (d_1|H)\ P\ (d_2|H)\ P\ (d_3|H) \qquad (3)$$

which states that the subjective posterior probability of a class given three cues is proportional to the product of the subjective probabilities for each cue given the class. Substituting from equation (2) into equation (3) yields:

$$P\ (H|d_1, d_2, d_3) \propto P\ (H|d_1)\ P\ (H|d_2)\ P\ (H|d_3) \qquad (4)$$

which states that the subjective probability of a class given three cues is proportional to the product of the subjective probabilities of the class for each cue. Equation (4) shows the form of Bayes's Theorem that was used to combine the probabilities in Beach's study.

In experiment (1) by Beach, twelve paid university students were run under three experimental conditions differing in complexity only. Capitol letters, numbers, and small letters served as cues and different colors made up the classes. Subjects were shown cards containing a combination of one cue from each of the cue dimensions, i.e., a capitol letter, a number, and a small letter, and the assigned class on the back. The cues were probabilistic indicants of the classes to which the card belonged. To construct the set of cards, Beach used a method which involved using a partial set of the larger possible set of cues. While this method does not produce independent cue dimensions, he justified its use on the basis of the limited experience the subjects were given with the cards. The method used for experiment (1) will not be explained; suffice it to say that Beach changed his procedure for experiments (2), (3), and (4) to include all possible sets of cues and that this revised approach was used in the present study and is explained in the chapter on Method.

Beach stated that the accurate knowledge of the probabilistic structure of the deck of cards would constrain subjects to be consistent. It is for this reason that the experiment was designed to reduce accuracy.

The first experimental condition consisted of a deck of 60 cards, 30 in each of two classes; the second consisted of 90 cards, 30 in each of three classes; and the third consisted of 180 cards, 30 in each of six classes. The subjects were told that all classes possessed an equal number of cards and that presentation order was random. They were then shown the cards one at a time and asked to make judgments about the class membership of the card on the basis of the cues shown. After the choice had been made, the correct class was identified and the experimenter proceeded on to the next card in the deck. This procedure was followed throughout the two presentations of the card decks for each of the three experimental conditions. These training trials were simply to provide the subjects experience with the probability based relationships of the experimental universe.

After training, the subjects were tested on cards possessing selected single cues and pairs of these same cues. For the two-class condition, subjects were shown eight single-cue and three paired-cue cards; the three class condition involved nine single-cue and five paired-cue cards; and the six-class condition had twelve single-cue and eight paired-cue cards. The single and paired-cue test cards were intermixed. The presentation was the same as during the training trials except that the correct class was not revealed to the subjects. The subjects were asked to estimate the probability of each card belonging to each of the possible classes by sliding markers along a 25-inch unmarked metal bar. All three classes were estimated on one bar at the same time so that the probabilities would add to one. Subjects were asked to make a second set of probability estimates for five of the test cards in order to obtain a measure of reliability. Subjects were run individually and for each experimental condition the five having the highest reliability were used in the analysis.

By inserting the subject's subjective probability estimates for two single cues given a particular class into equation (4), calculating the product and correlating the obtained values with the subject's paired-cue estimate for the same class, Beach obtained a measure of revision consistency for each subject. To evaluate paired cue accuracy, the objective training deck probabilities were substituted into equation (4) and the resulting posterior probabilities correlated with the paired-cue subjective estimates. Single-cue accuracy was determined by correlating the objective single-cue probabilities with the subject's subjective estimates.

The revision consistency analysis for experiment (1) resulted in what Beach considered positive support for the Bayesian model. Beach eliminated seven of the twelve subjects by reporting the consistency results on only the five who obtained the highest reliability coefficients. This was done for each of the three experimental conditions. Whether or not the same five subjects obtained the highest reliability correlations under each condition, was not reported. The results of experiment (1), (Table I), show the consistency and reliability correlations for each experimental condition. Of the 15 consistency correlations, 13 are significant at the .05 level and above. Beach suggested that the lower revision consistency results for the six-class condition probably resulted from the increased confusion and opportunities for error when estimating probabilities for six-classes. The single-cue accuracy correlations were, in terms of mean correlations, .46, .31, and ,34 for the two-, three-, and six-class conditions. The mean correlations for paired-cue accuracy were .04, .06, and .14 for the three experimental conditions.

Experiment (2) by Beach (1966), directly investigated the hypothesis that fluctuations in the accuracy of subjective probability estimates should not influence revision consistency. This hypothesis was examined when accuracy fluctuated as a result of forgetting the probabilities and to additional training. The training cards from the three-class condition were used. The training procedure was the same except that subjects went through the training deck four times instead of twice in order to permit some degree

TABLE I

RELIABILITY AND CONSISTENCY CORRELATIONS FOR EACH EXPERIMENTAL CONDITION

Experimental Condition	Subjects	Reliability Correlations	Consistency Correlations
Two-class	1	.99	.94
	2	.97	.99
	3	.94	.93
	4	.92	.70
	5	.92	.95
	$\bar{X}$	.96	.94
Three-class	1	.92	.94
	2	.87	.77
	3	.86	.94
	4	.97	.99
	5	.92	.95
	$\bar{X}$	.92	.94
Four-class	1	.87	.42
	2	.68	.68
	3	.92	.74
	4	.80	.42
	5	.72	.69
	$\bar{X}$	.82	.59

of accuracy. The subjects were 30 paid men university students trained and tested in groups of about five at a time. The subjects were divided into a short delay group and a long delay group. Both groups were tested, using the same procedure as was used in Beach's first experiment. The short delay group returned one or two days later and the long delay group returned five days later. When they returned, both groups were tested again, given one training trial, and tested a third time. Although this procedure did produce a decrease in accuracy for the long delay group that was statistically significant, it was not as much as expected, and 15 subjects, some from each group, were asked to return two months later to take a fourth test. This attempt was also unsuccessful in that the resulting decrease in accuracy was not statistically significant.

Beach's results (Table II) were reported in terms of mean correlations. It is interesting to note that the paired-cue accuracy correlations actually increased over the two- and five-day delay periods. Beach does not comment on this increase and simply states that paired-cue accuracy remained low throughout the experiment. The fact that the two-month delay resulted in less decrease in accuracy than the five-day delay is also surprising. The mean consistency correlations do not appear to be influenced by the only significant decrease in accuracy found over the five-day delay period.

TABLE II

SINGLE- AND PAIRED-CUE ACCURACY AND CONSISTENCY
CORRELATIONS FOR EACH DELAY PERIOD

	Mean Correlations for Single-Cue Accuracy			Mean Correlations for Paired-Cue Accuracy			Mean Correlations for Consistency		
Test	I	II	III	I	II	III	I	II	III
Short delay 1 - 2 days	.49	.45*	.60	.19	.36	.38	.86	.87	.86
Long delay 5 days	.53**	.38***	.52	.28	.35	.45	.76	.79	.87
Test	III	IV		III	IV		III	IV	
Two-month delay	.54	.43		.44	.34		.94	.90	

* P .005
** P .01
*** P .05

Beach correlated the changes from one test session to another in the subjects' reliability, accuracy, and consistency correlations. The results of this analysis showed that the changes in single- and paired-cue accuracy from one test session to another were not related to increases or decreases either in consistency or in reliability. He did find, however, a significant positive correlation between changes in consistency and changes in reliability. The results, according to Beach, supported the notion that reliability places an upper limit on revision consistency correlations but does not dictate high consistency. He found that at high levels of reliability there were fairly large individual differences in consistency and that at low levels of reliability the results were biased toward low consistency.

Experiment (3) by Beach (1966) had two objectives. The first was to examine the subjects' ability to revise displayed deck probabilities and the second was to see if the subjects' class choices for cards during training would correspond to their subjective probability estimates for these same cards.

The revision of subjective probabilities was also examined in the same manner as experiment (1). A three-class, two-dimensional universe was used and the subjects were 32 male students who were paid for their participation.

The available results show that for the 16 reliable subjects the mean single-cue accuracy correlation was .43. The mean paired-cue accuracy correlation was .31 and the mean revision consistency correlation was .79.

To examine the subjects' ability to revise deck (objective) probabilities, the subjects were shown bar graphs of the actual cue-class relationships for single cues and were asked to make estimates of the probabilities for a card that possessed pairs of such cues. The results showed that 13 out of 32 subjects simply averaged the probabilities across the two cues for each class. The 19 subjects who did not average obtained a mean correlation between their revised paired-cue estimate and the Bayesian revisions of the deck probabilities of .96.

The data for examining the relationship between class choices and paired-cue probability estimates were obtained from test cards interspersed in the fourth and last training deck. Sixteen reliable subjects made a class choice for each of four test cards during the fourth training deck trial and later were tested with these same cards and asked to make paired-cue estimates. The mean single- and paired-cue accuracy correlations for the 16 reliable subjects were .43 and .31. Eighty per cent of the class choices corresponded to the class later given the highest subjective probability estimate. The results further showed that 7 of the

16 subjects had all four choices correspond to their high estimates; for five subjects, three choices corresponded; and for four subjects, only two corresponded. These findings, according to Beach, appear to support the notion that even though a subject's subjective probabilities are quite inaccurate, they are still utilized for decisions about a card's class membership.

Experiment (4) by Beach (1966) examined whether or not subjects could consistently revise estimates for pairs of cues when only single-cues had occurred on the cards throughout training.

Revision consistency in this situation was examined under the same procedural model as experiment (1), with the deck being presented six times. The subjects were tested by having them make bar graphs of their estimates for seven single-cue test cards and for six paired-cue test cards. Analyzing the data for the 8 reliable subjects (there were 10 in the group), Beach obtained the following results: the mean single-cue accuracy correlation was .44; the mean paired-cue accuracy correlation was .50; and the mean revision consistency correlation was .76.

In the discussion of all four experiments, Beach concludes that the overall results support the notion that "subjects possess a rule for revising subjective probabilities which they apply to whatever subjective probabilities they have at the moment; whether the probabilities are derived from experienced relative frequencies or from displayed probabilities, whether they are accurate or inaccurate, whether they have just been learned or whether they are the residuals of partially-forgotten values learned days before. Through all of these variations in the accuracy of their subjective probabilities the subjects' rule remains unchanged." Beach suggested that the revision rule had been shown to be essentially Bayes's Theorem.

In summary, Beach's support for the Bayesian model of revision consistency consists of three groups of subjects with high average consistency correlations and a minimum of five subjects who individually obtained significant consistency coefficients. In each experiment only a certain number of subjects obtaining the highest reliability correlations were used in the analysis of consistency. The revision consistency results of experiments (2), (3), and (4) are reported in terms of mean correlations ranging from .76 to .87. The results of experiment (1) show consistency correlations on five of 12 subjects for each experimental condition. Because the possibility exists that the same five subjects were used for the consistency analysis under each condition, the results, while impressive, may be applicable to a sample of only five subjects.

The present study was designed to investigate the Bayesian revision consistency results of Beach (1966). While the investigation is not a replication of Beach, it does follow the general model of experiment (1) and therefore can be identified as an attempt to determine the extent to which Beach's results can be generalized. Specifically, the hypotheses were as follows: (1) Subjective probability estimates for combined-cue stimuli are highly consistent Bayesian combinations of the subjective estimates for the corresponding single-cue stimuli; (2) This revision consistency will be high even though the subjects' accuracy is low when compared to the objective probabilities.

There are several of Beach's experimental conditions which may have biased his studies in such a way as to restrict the generalization of his results. Questionable conditions include: (1) Beach utilized cues which probably did not form a "whole" when seen together. Even though the subjects were told to infer the card's class membership, it seems likely that they were seeing three individual cues and not the card as an object; (2) Beach had subjects estimate subjective probabilities in such a way as to insure that they would add to one. It seems likely that by fixing the subjective probabilities on the last class judgment for each cue, a certain amount of consistency would be expected, especially for the two- and three-class conditions which provided the significant support for the Bayesian revision rule; (3) Beach examined consistency for the Bayesian revision of pairs of cues only. That is, he combined only the minimum number of cues using Bayes's Theorem.

In part, the design of the present study was an attempt to modify the questionable conditions of Beach's experiments. In this way, it was thought that the generalizations about the Bayesian revision rule could be validated. In general, the present design follows the model of experiment (1) by Beach (1966) with the following modifications: (1) the experimental universe was changed in an attempt to have the cues when seen together form an object; (2) subjective probability estimates were made by having subjects indicate number values along separate continuums for each class and the additivity of the values was left to their discretion; (3) revision consistency was examined for combinations of three cues.

Beach utilized the form of Bayes's Theorem as shown by equation (4) on the basis that the subjects believed that the classes were equally probable. The basis for his study rests on this specific Bayesian equation, and his method of fitting events to the model. The present study consisted of two experiments designed to examine this position. It is possible that even though the subjects in Beach's study were told at the beginning of the experiment that the classes were equally represented, they might not have been reacting accordingly during the test trials. In experiment (1) of the present study, an attempt to eliminate this possibility required subjects to report their subjective opinions regarding the class distribution, after the experiment. Because of the relatively simple (four-class) universe, it was hypothesized that most of the subjects would report equal class proportions. The subjects that reported equal class probabilities were to form the Bayesian analysis group. It was thought that if a sufficient number of subjects indicated that the classes were not equally represented, it would be profitable to compare their Bayesian revision consistency results obtained "illegally" with the Bayesian analysis group, who fit the model. Theoretically, the Bayesian analysis group should do better. Experiment (2) of the present study incorporated Beach's method of advising the subjects about the equal class representation.

CHAPTER II

METHOD

Subjects

The subjects used for experiment (1) of this study consisted of 43 college undergraduates enrolled spring quarter, 1966, at Central Washington State College. The subjects used for experiment (2) consisted of 14 students from the same population. All of the students participating in the experiment were satisfying a class requirement.

Apparatus

The experimental Universe consisted of sixty-three 5 x 8 index cards. Each card had a cue value from each of three different multi-valued cue dimensions on its face and one of three possible class labels on its back. The three cues when seen together formed an object identified as a spaceman. The cue dimensions were: (1) the geometrical figures used for the body of the spaceman, (2) the colors of the body, (3) the shapes of the legs. Each dimension consisted of four individual cues which were: (1) the body shapes of square, circle, cross, and triangle; (2) the colors of red, yellow, blue, and green; (3) the leg shapes of dashed lines, wavy lines, straight lines with cross marks, and continuous lines which slanted out and curved up at the bottom. The classes to which the spacemen were assigned consisted of the planets of Mars, Pluto, and Moon. (Refer to Appendix I for example drawings of the spacemen.)

Dr. Beach, in personal correspondence dated March 1966, provided a detailed explanation of his revised method for arriving at the cue combinations and the assignment of cues to classes. Beach's method, which was used for this study, involves the use of all possible sets of cues and assures independent cue dimensions. A three-dimension (4x4x4) matrix for the cue dimensions, body shape, color, and leg shape was used. Each cell of the matrix was randomly assigned to one of the three classes in such a way that each class had an equal number of cells. The proportion of cells belonging to each class was then calculated for each row and column. The proportions were the conditional probabilities of each class being correct for a card possessing the cue associated with the row and column and constituted the objective deck probabilities. Because the combinations formed in this manner resulted in too many infirming cases, the class assignments were adjusted, which was in accordance with Beach's method. Infirming cases are described by Beach (1964) as being special cue combinations for which the relative frequency of cue-to-class association values lead to a class choice which is not the one assigned and designated as correct. Appendix II shows the cue combinations and class assignments that comprised the experimental

Universe. Table III shows the relative frequency breakdown of cue-to-class assignments.

The entire deck of 63 cards was used during the training trials. The cards used for the test trials consisted of nine cards possessing single cues, three from each dimension, and three cards possessing combinations of three of the single cues. When seen in combination, the cues formed a spaceman. Because the training trial deck consisted of all possible combinations of cues (less one so that the classes had an equal number of cards), the three-cue combination test cards were obtained directly from the training deck. Appendix III shows the test trial cues and cue combinations in order of presentation. Five of the test cards were presented to the subjects a second time in order to obtain a measure of reliability.

Subjects' responses were recorded on preprinted answer sheets. The response forms for the training trial consisted of the three classes, Mars, Pluto, and Moon, printed together and each group numbered from 1 to 126. The test trial forms consisted of the three classes printed together and numbered from 1 to 17. Under each number and extending across the page from each class was a straight-line continuum with a "0" on the extreme left and "100" on the extreme right. (Refer to Appendix IV for the response form.)

TABLE III

FREQUENCY DISTRIBUTIONS FOR CUE VALUES ASSOCIATED WITH EACH CLASS AND THE CORRESPONDING PROBABILITIES OF THE CLASSES GIVEN EACH CUE

CLASS	BODY SHAPE			
	Circle	Cross	Triangle	Square
MARS: f	3	2	7	9
P(Class/cue)	.19	.12	.44	.60
PLUTO: f	8	10	1	2
P(Class/cue)	.50	.63	.06	.13
MOON: f	5	4	8	4
P(Class/cue)	.31	.25	.50	.27
	COLOR			
	Red	Yellow	Blue	Green
MARS: f	9	2	4	6
P(Class/cue)	.56	.12	.25	.40
PLUTO: f	3	3	9	6
P(Class/cue)	.19	.19	.56	.40
MOON: f	4	11	3	3
P(Class/cue)	.25	.69	.19	.20
	LEG SHAPE			
	Dashed Lines	Wavy Lines	Lines with crossmarks	Lines which curve up
MARS: f	5	13	1	2
P(Class/cue)	.31	.82	.06	.13
PLUTO: f	10	1	4	6
P(Class/cue)	.63	.06	.25	.40
MOON: f	1	2	11	7
P(Class/cue)	.06	.12	.69	.47

Procedure

The 43 subjects in experiment (1) met in groups of 10 and 11. They were tested in their classroom during the regular class hour. The experimenter and an assistant sat side by side at a table in front of the room. A 10 x 36 inch piece of white cardboard was placed upright on the table to shield irrelevant activity from the subjects' view. When the subjects were seated, the experimenter held up a blank 5 x 8 card and asked that they move to the front and sit in a position so as to be able to see the card clearly. When the subjects were settled and the answer sheets handed out, the experimenter read the training trial instructions. The instructions consisted of a general introduction to the experiment and an explanation of the task and experimental procedures. Subjects were told that the cards contained the information that they would need to tell where a particular spaceman came from, i.e., Mars, Pluto, or Moon. The task was to learn to use the information. Subjects were told that after seeing a spaceman, they were to classify him according to his planet by marking the preprinted answer sheet. It was pointed out that in the beginning they would be guessing, but that they were to make a choice regardless. (Refer to Appendix V for the specific instructions.) The experimenter responded to questions and the training trials began.

The cards were shuffled by the assistant and handed one at a time to the experimenter. While looking at a stop watch, the experimenter held each card above the cardboard shield for five seconds; the card was then removed from view for three seconds; the subjects were instructed to circle their answer; and the card was raised again long enough for the experimenter to identify the correct class. After a card had been presented, the experimenter handed it back to the assistant and obtained the next one. In this way the single deck was again shuffled and the experimenter was able to present the same deck of 63 cards twice.

After the training trials, the experimenter read the instructions for the test trials. The subjects were told that they would not be able to see all of the spacemen on some of the cards. They were asked to estimate the probability that the information shown, i.e., part of, or the whole spaceman, indicated that a spaceman was from each planet. It was explained that the probability estimates for each class would be made by placing marks and number values (from zero to one hundred) along a continuum. A basic definition of probability was included in the instructions. (Refer to Appendix VI for the specific test instructions.) The subjects were given a sample trial and questions were answered. With the presentation of a card the experimenter said, for example, "Estimate the probability that the body shape 'square' indicates that a spaceman is from Mars, Pluto, and Moon." Each card was held in view of the subjects for 20 seconds. After the experiment, the subjects were given a questionnaire and asked to indicate their impression of the approximate number of spacemen from each planet. (Refer to Appendix VII for the questionnaire.) Due to an error in the experimental procedure, the subjects did not complete the questionnaire until six days after the experiment. A criticism of this procedure would concern subject recall. The questionnaire allowed the subjects to respond by stating that they could not remember, did not have an idea at the time, or that they believed the classes to be equally or unequally represented. It was thought that those subjects who indicated equal class proportions must have been under that impression at the time of the experiment.

The procedure for experiment (2) was the same as experiment (1) except that the subjects were told in the training trial instructions that they would see an equal number of spacemen from each planet. The subjects for the second experiment were not given the post-experiment questionnaire. The 14 subjects in experiment (2) were tested as a group during their class hour in their regular classroom.

CHAPTER III

RESULTS

The purpose of these experiments was to examine whether subjects' pre- and post-revision subjective probabilities are highly consistent in the manner described by Bayes's Theorem. The model of experiment (1) by Beach (1966) was followed in that subjects were allowed only minimal experience with the probability based cue-class universe. In other words, revision consistency was examined in a situation in which over-all accuracy was expected to be low when compared with the objective probabilities.

Four types of analyses were utilized:

1. Reliability was evaluated by obtaining Pearson product moment correlations between the subject's first and second estimates for the five reliability cards.
2. Single-cue accuracy was examined by obtaining Pearson product moment correlations between the subject's single-cue estimates and the corresponding objective probabilities.
3. Combined-cue accuracy was evaluated by obtaining Pearson product moment correlations between the subject's combined-cue card estimates and the Bayesian predictions for the corresponding objective, single-cue probabilities.

4. Revision consistency was examined by obtaining Pearson produce moment correlations between the subjects' combined-cue card estimates and the Bayesian predictions based on their subjective estimates of the corresponding single-cue cards.

Out of the 43 subjects in experiment (1), 19 indicated that there was an equal number of spacemen in each class. These 19 subjects made up the Bayesian analysis group. The results, as shown in Table IV, revealed that only eight out of the 19 subjects obtained revision consistency correlations which were significant from zero. (Correlational significance when not otherwise stated will always refer to significance from zero with alpha set at .05.) There were 11 subjects who obtained reliability correlations of .80 and higher (the .80 cut-off is arbitrary); only five of these high reliable subjects obtained revision consistency correlations that were significant. A Chi square test of significance on the number of high reliable subjects who obtained significant and non-significant consistency correlations showed that a null hypothesis of a 50:50 proportion could not be rejected. The mean revision consistency correlation for all of the 19 subjects was .53. The mean consistency correlation for the 11 high reliable subjects was .55.

The 19 subjects in the Bayesian analysis group obtained a mean reliability correlation of .73. Out of the 19

TABLE IV

RELIABILITY, ACCURACY, AND CONSISTENCY CORRELATIONS FOR EACH SUBJECT IN THE BAYESIAN ANALYSIS GROUP

Subjects	Reliability	Single-cue Accuracy	Combined-cue Accuracy	Consistency
No. of Estimates	15	27	9	9
1	.88**	.34*	.32	.94**
2	1.00**	.60**	.90**	.99**
3	.91**	.33*	.22	.35
4	.96**	.46**	-.32	.36
5	.90**	.41*	.65*	.57
6	.54*	.17	.36	.51
7	.86**	.53**	.59*	.68*
8	.81**	.02	-.001	.41
9	.74**	.37*	.63*	.61*
10	.94**	.41*	.50	.10
11	.96**	.40*	.10	.17
12	.76**	.25	-.22	.81**
13	.85**	.43*	.45	.83**
14	.47*	.25	-.02	.34
15	.95**	.09	-.13	.64*
16	.47*	.43*	.45	.78**
17	.42	.44*	.90**	.30
18	.21	.20	.91**	.50
19	.21	.60**	.72*	.25
$\overline{X}$	.73	.35	.37	.53

*P $\leq$ .05
**P $\leq$ .01

reliability correlations, 16 were significant at the .05 level and above. The mean single-cue accuracy correlation was .35. Although 13 of the 19 single-cue accuracy correlations were significant, their magnitudes tended to be quite low. The mean combined-cue accuracy correlation was .37. In this case, 7 out of the 19 were significant.

The results for the non-Bayesian analysis group are shown in Table V. Out of the 43 subjects in the experiment, 18 indicated that they believed that there was not an equal number of spacemen in each class. Out of the 18 subjects, 8 obtained consistency correlations that were significant at the .05 level and above. There were 10 out of the 18 subjects who attained reliability correlations greater than .60. Five of the 10 high reliable subjects obtained significant consistency correlations. The mean revision consistency correlation for all of the subjects was .50; for the high reliable subjects it was .53.

Significant reliability correlations were obtained by 13 out of the 18 subjects in the non-Bayesian analysis group. The mean reliability correlation was .59. Although eight subjects obtained significant single-cue accuracy correlations, they tended to be quite low, which is reflected by the mean correlation of .28. Significant combined-cue accuracy corellations were obtained by 8 out of the 18 subjects. The mean combined-cue accuracy correlation was .47.

TABLE V

RELIABILITY, ACCURACY, AND CONSISTENCY CORRELATIONS FOR EACH SUBJECT OF THE NON-BAYESIAN ANALYSIS GROUP

Subjects	Reliability	Single-cue Accuracy	Combined-cue Accuracy	Consistency
No. of Estimates	15	27	9	9
1	.31	.30	.69*	.70*
2	.34	.36*	-.13	.43
3	.53*	.10	-.07	.73*
4	.90**	.40*	.26	.81**
5	.54*	.27	.91**	-.27
6	.62**	.19	.32	.35
7	.17	-.13	-.20	.46
8	.51*	.57**	.50	.56
9	.88**	.07	.71*	.14
10	.33	.65**	.92**	.80**
11	.31	.33*	.19	.28
12	.62**	.30	.32	.46
13	.95**	.25	.83**	.77**
14	.67**	.48**	.61*	.79**
15	.62**	.09	.91**	.11
16	.91**	.07	.36	.48
17	.75**	.37*	.39	.66*
18	.68**	.41*	.93**	.75**
$\bar{X}$	.59	.28	.47	.50

*P ≤ .05
**P ≤ .01

Six subjects indicated that they had no idea about the representation of each class. These subjects were not included in the analysis of consistency.

The results for experiment (2) (Table VI) appear to follow the same general trend as experiment (1). Out of the 14 subjects who were told that the classes were equally represented, 5 obtained positive revision consistency correlations that were significant at the .05 level and above. There were 7 subjects who obtained reliability correlations of .60 and greater. Of these 7, only 2 obtained significant consistency correlations. The mean consistency correlation for the entire group was .30. The mean consistency correlation for the high reliable subjects was .26.

Out of the 14 subjects in experiment (2), nine obtained significant reliability correlations. The mean reliability correlation was .52. Ten subjects obtained single-cue accuracy correlations that were significant. These correlations tended to be relatively low with a mean coefficient of .35. There were four subjects who obtained significant combined-cue accuracy correlations. The mean coefficient in this case was .36.

It is interesting to note that out of the 57 subjects used in this study, 41 obtained reliability correlations that were significant. When considered in light of the considerably lower single-cue accuracy correlations, these

TABLE VI

RELIABILITY, ACCURACY, AND CONSISTENCY CORRELATIONS
FOR EACH SUBJECT IN EXPERIMENT (2)

Subjects	Reliability	Single-cue Accuracy	Combined-cue Accuracy	Consistency
No. of Estimates	15	27	9	9
1	-.05	.42*	.54	.67*
2	.77**	.42*	-.54	-.59*
3	.80**	.18	.63*	-.12
4	.28	.38*	.37	.67*
5	.70**	.34*	-.29	.20
6	.47*	.13	.19	.54
7	.42	-.07	-.10	-.11
8	.50*	.53**	.55	.79**
9	.91**	.48**	.44	.43
10	.64**	.39*	.46	.22
11	.64**	.34*	.82**	.48
12	.77**	.66**	.86**	.97**
13	.71**	.55**	.75**	.72*
14	.18	.21	.37	-.65*
$\bar{X}$	.52	.35	.36	.30

*P $\leq$ .05
**P $\leq$ .01

results seem to support the notion that subjects will utilize and stick with a subjective probability judgment even though it is quite inaccurate.

The responses for the 33 high reliable subjects in experiment (1) were grouped and evaluated in terms of a gross indication of direction. It was noticed that the class having the highest average probability estimate for the combined-cue card distributions corresponded to the class having the highest average estimate for the related single-cue test card distributions. In other words, supposing test card number eight consisted of a spaceman having a square shape, the color red, and dashed lines for legs. By adding the subjects' estimates for each class, i.e., Mars, Pluto, and Moon, across subjects and determining the mean of the distribution for each class, it was possible to obtain a gross indication of which class was considered the "best bet" by the group. By following the same procedure for the corresponding single-cue test cards, i.e., three cards that possessed one of either the square shape, the color red, or dashed lines for leg shape, for example, it was possible to obtain another gross indication of what the group considered to be the "best bet." It was found that the "best bet" for the combined-cue test cards corresponded to the "best bet" for the related single-cue test cards. The results are shown in Table VII.

TABLE VII

AVERAGE OF 33 SUBJECTS' ESTIMATES FOR THE COMBINED AND RELATED SINGLE CUE TEST CARDS FOR EACH CLASS

Combined-cue Test Card Number	Class Means (probabilities)			Related single-cue Test Card Numbers	Class Means (probabilities)		
	Mars	Pluto	Moon		Mars	Pluto	Moon
4	.30	.37	.39*	2,9,10	.33	.39	.39*
8	.63*	.25	.20	3,6,11	.47*	.31	.35
12	.32	.20	.57*	1,5,7	.35	.30	.45*

* "best bets"

It appears possible to predict which class (Planet) would be considered the most likely choice for a particular spaceman by simply adding the subjects' probability estimates for the cues which comprised the spaceman and selecting the class with the highest value. These gross indications suggested that a simple additive model might account for the manner in which subjects combine their estimates for single cues. An analysis (Table VIII) correlating the subjects' subjective probability estimates for the combined-cue test cards with the average estimate of the related single-cue cards revealed that 17 out of the 33 high reliable subjects in experiment (1) obtained consistency correlations that were significant at the .05 level and higher. The mean consistency correlation for the group was .63.

The Bayesian revision model accounted for 8 of 19 subjects in the Bayesian analysis group with a mean consistency correlation of .53, and 5 of the 14 subjects in experiment (2) having a mean consistency coefficient of .30. The additive model accounted for 17 of 33 subjects having a mean consistency correlation of .63.

TABLE VIII

REVISION CONSISTENCY CORRELATIONS FOR 33 RELIABLE SUBJECTS USING THE ADDITIVE MODEL

Subjects	Reliability	Consistency
Number of Estimates	15	9
1	.88	.87**
2	1.00	.99**
3	.88	.27
4	.91	.38
5	.96	.57
6	.90	.77**
7	.54	.39
8	.86	.72*
9	.53	.76**
10	.90	.87**
11	.54	-.38
12	.62	.16
13	.70	.10
14	.81	.41
15	.51	.52
16	.88	.25
17	.74	.66*
18	.94	.39
19	.71	.60*
20	.96	.47
21	.62	.65*
22	.95	.75**
23	.73	.80**
24	.73	.67*
25	.67	.79**
26	.76	.76**
27	.62	.34
28	.92	.35
29	.85	.83**
30	.76	.77**
31	.68	.77**
32	.76	.18
33	.95	.57
$\bar{X}$		.63

*$P \leq .05$
**$P \leq .01$

CHAPTER IV

DISCUSSION

The purpose of these studies was to provide a basis for generalizing the results of an experiment (1) by Beach (1966). The relationship between subjects' subjective probability estimates for combinations of cues and estimates for the single cues which comprised the combinations formed the basis for the study. Beach suggests, on the basis of his results, that this relationship is essentially Bayes's Theorem.

In order to validate and provide evidence upon which to further generalize the application of Bayes's Theorem in this type of experimental setting, it is necessary to define more specifically the results that can be interpreted as being in support of the position. When Bayes's Theorem is hypothesized to be the appropriate normative model for how opinions are revised, it must be assumed to be applicable to all persons. In other words, it is posited that <u>all</u> men are intuitive staticians and have in common a Bayesian methed of information processing. This is not to say that experimentally every subject must be highly consistent; certainly individual differences in attitude and interest in the experiment, for example, would allow for a few subjects to function in ways other than would be predicted by Bayes's

Theorem. In utilizing the technique of correlation to determine the degree of Bayesian consistency, the level of the correlation coefficient that is to be labled "highly consistent" must be determined. Because there is no evidence upon which to base a specified value of the population correlation, "highly consistent" should meet the minimal requirement of being significantly greater than zero. In summary, experimental evidence which is to be interpreted as being in support of the Bayesian revision rule must consist of most subjects obtaining consistency correlations that are significant.

The present investigation follows the general model of Beach's study in that subjects were allowed only sufficient experience with the experimental universe to develop some inaccurate notions about the probabilistic rules which linked cues to classes. The results show that this was accomplished in that the correlations between the subjects' estimates and the objective probabilities tended to be quite low. The mean single-cue accuracy correlations were .35, .47 and .36 for the three experimental groups. A comparison of these findings with the accuracy results of Beach is of no concern since the rational for having subjects attain only a minimal amount of accuracy was based on Beach's position that high accuracy would constrain subjects to be consistent. In other words, as long as the accuracy is not high across subjects, the model of Beach's study is upheld. The single-cue accuracy for the present study, while slightly lower across subjects, compares quite favorably with the results of Beach. The combined-cue accuracy attained by the subjects in the present study was somewhat higher than in Beach's investigation.

The use of a reliability control in the present study was necessary in order to identify subjects who were simply writing down numbers at random. An examination of the protocols for the 16 subjects in experiment (1) who obtained reliability correlations of .80 and higher, revealed that the test-retest measurement involved something more than the subjects simply remembering earlier response. Only one subject had stereotyped all five of his responses; another stereotyped four responses; four subjects made identical estimates on two reliability cards and one subject on one card. The remaining nine subjects gave responses which were linearly related but which were not alike. These observations only add support to the notion that subjects will utilize a subjective judgment, even though inaccurate, and stick with it, without, in most cases, simply remembering their earlier response. Certainly, stereotyped responses cannot be considered a weakness of the measurement in that the subjects may have been honestly expressing their judgment rather than trying to outguess the experiment. As was previously stated, 41 out of the 57 subjects participating in the experiments attained reliability coefficients that were significant. When considered in light of the low single-cue accuracy correlations, these results also seem to support the notion that subjects will utilize and stick with an inaccurate subjective probability judgment. Although it was necessary to define significant reliability in terms of a zero correlation, the results have been written to include only the highest obtained reliability coefficients for the experimental groups. This was done to more closely approximate the reliabilities associated with the Bayesian consistency results of Beach.

Two experiments were run to test the hypothesis of the Bayesian revision rule. As in Beach's study, the experimental procedures were designed to fit the requirements of the particular form of Bayes's Theorem that was utilized. The equation is:

$$P'(H/d_1, d_2, d_3) \propto P\ (H/d_1)\ P\ (H/d_2)\ P\ (H/d_3)$$

which states that the subjective probability of a class given three cues is proportional to the product of the subjective probabilities of the class given each cue. In order to justify using this equation, it is necessary that the subjects consider the classes to be equally represented and therefore having an equal probability of occurring in the deck. It is only in this way that the probability of a class can be removed from the basic equation. Experiment (1) of the present study assured that the subjects believed that the classes were equally probable by including a procedure in which the subjects were asked after the experiment to state their opinion. Those 19 subjects fitting the model of the Bayesian equation were labled the Bayesian analysis group. There were 18 subjects who indicated that the classes were not equally represented and these subjects constituted the non-Bayesian group. Experiment (2) consisted of 14 subjects who were told at the beginning of the experiment that they would see an equal number of cards from each class.

The results for the Bayesian analysis group do not meet the requirements for support of the Bayesian revision rule. Out of the 11 subjects who obtained reliability correlations of .80 and higher, only 5 had consistency correlations that were significant from zero. A Chi square test showed that a null hypothesis of a 50:50 proportion could not be rejected. These results are distant enough from the "all-men" requirement of a normative model to be considered negative.

The results of experiment (2) do not support the hypothesis of the Bayesian revision rule. Of the seven subjects who obtained reliability correlations of .60 and higher, only two had consistency correlations that were significant.

In conclusion, the results do not support the hypothesis. Subjective probabilities for the combined-cue stimuli cannot be said to be highly consistent Bayesian combinations of the subjective estimates for the corresponding single-cue stimuli.

The present experiments differed from the experiment by Beach (1966) in several significant ways: (1) a different type of cue-class universe was used; (2) subjective probabilities were not forced to add to unity; and (3) revision consistency was examined for combinations of three cues rather than two.

Bayes's Theorem is based on the fact that probabilities add up to one. If Bayes's Theorem was used as an equality, this condition would have to be met. In the form of a proportion, however, with the results defined in terms of correlation, the relationships will not be influenced by subjective probabilities which when summed exceed or are less than unity. Modifications (2) and (3) further complicate the situation in an attempt to provide a basis for generalization, but should theoretically have no effect on the outcome. It could be argued, however, that the change in the experimental universe (from letters and numbers to spacemen) might account for the differences in the results.

A criticism could be made of the experimental universe based upon the attention paid to the various cues that comprised the spaceman. While one would expect subjects to pay approximately equal attention to the rather homogeneous cues of capital letters, small letters, and numbers, the expectancy might not be so high when using such diverse cues as geometrical shapes, colors, and various leg shapes. If this were the case, it would be possible for subjects to attain high reliability correlations, but the reliance on one or two cues would not result in high consistency as determined by the Bayesian equation. The present experiment does not permit a direct examination of this question. It would seem, however, that if the cues possessed varying degrees of attractiveness, a significant difference in the variances of subjective probability estimates across cue dimensions would ensue. For example, if colors are relied upon to a greater extent than leg shapes, then it would be expected that the variance between the estimates for the cue dimension leg shapes would be significantly greater than the variance for the color dimension. An examination of the variances for the various cue dimensions revealed that such a difference did not exist. The data were obtained from the 43 subjects in experiment (1). The results are shown in Table IX.

It can be seen that there is greater intra-dimensional variance than between cue-dimensions. It seems plausible, therefore, that the discrepancy between the results of the

TABLE IX

STANDARD DEVIATIONS OF SUBJECTIVE PROBABILITIES FOR THE THREE CUE DIMENSIONS

Cue Dimensions	Mars	Pluto	Moon
<u>Colors</u>			
Red	.23	.19	.15
Green	.23	.24	.22
Yellow	.24	.21	.23
$\overline{X}$ = .22			
<u>Leg Shapes</u>			
Type I	.23	.24	.22
Type II	.27	.20	.21
Tupe IV	.23	.21	.25
$\overline{X}$ = .23			
<u>Body Shapes</u>			
Square	.23	.20	.23
Circle	.21	.18	.21
Triangle	.25	.23	.25
$\overline{X}$ = .22			

present study and Beach's might be attributed to the limiting factors in Beach's study.

The consistency results of the non-Bayesian group compared quite favorably with the Bayesian analysis group. Even though the reliability correlations tended to be lower (mean correlation of .59 as compared to .73 for the Bayesian group), out of the ten subjects who obtained reliability coefficients of .60 and higher, five had consistency correlations that were significant. The mean, high reliable, consistency correlation for the non-Bayesian group was .53. Obviously the hypothesis of a 50:50 proportion is supported in this case also. At first consideration, these findings appeared to cast doubt on the form of Bayes's Theorem upon which the hypothesis rests. It seemed likely that not taking into consideration the differential class probabilities would constrain the results of the non-Bayesians toward less consistency. The fact that both groups obtained comparable consistency results appeared to place the specific equation, which required conditions that the non-Bayesians had not met, in serious question. If the equation is of doubtful applicability, then negative results would be meaningless in terms of the general Bayesian hypothesis. However, the tenability of the equation will be discussed in view of additional evidence.

The test card responses for 33 high reliable subjects in experiment (1) were grouped and evaluated. It was noticed that the class having the highest average probability estimate for the combined-cue cards, corresponded to the class having the highest average estimate for the related single cue test cards. These gross indications suggested that a simple averaging process might account for the manner in which subjects combine their estimates for single cues. The results of applying the additive model to the manner in which the 33 high reliable subjects in experiment (1) combined their estimates for single cues, show that the additive model can account for how subjects revise their opinions equally as well as the Bayesian equation. Of the 33 subjects whose single-cue responses were simply added and then correlated to their related combined-cue judgments, 17 attained consistency correlations that were significant. The mean consistency correlation was .63. These results could be interpreted to mean that it is not necessary to rely on the complicated Bayesian approach to information processing.

Upon examination of the 17 subjects who attained significant consistency correlations using the additive model, it was found that 14 of these subjects had also achieved significant consistency under the Bayesian model. Of the 14, six were from the Bayesian analysis group and eight were non-Bayesians. The fact that 14 out of 17 subjects gave responses which resulted in high consistency correlations for both the Bayesian and additive models shows very clearly the lack of sensitivity inherent in correlational analysis. It follows that rather than placing any high value on the results of the additive model, it may be best to question the use of correlation as the analytical technique used to identify the relationships.

Apparently the numerical values of the subjective probabilities for most of the subjects were such that high consistency correlations could result from either an averaging process or combining according to Bayes's Theorem. Table X provides an example of how this might work.

It can be seen that the same relative position of the classes is maintained using both models. For correlational analysis to distinguish between the Bayesian revision rule and the process of averaging would require the subjects to utilize more extreme or very low subjective probability values. Table XI provides an example of how this might work.

It can be seen that while class (2) is the "best bet" using the Bayesian model, the additive model predicts that the subjects will place the highest probability value on class (1) when the cues are seen together.

Because of the lack of sensitivity in the present correlational analysis, conclusions regarding the applicability

TABLE X

HYPOTHETICAL PROBABILITY ESTIMATES FOR CLASSES GIVEN SINGLE CUES

Class	Cue A	Cue B	Cue C	
Bayesian Model				
1	(.30)	(.60)	(.50)	.090
2	(.20)	(.10)	(.25)	.005
3	(.50)	(.30)	(.25)	.038
Additive Model				
1	.30+	.60+	.50	1.40
2	.20+	.10+	.25	.55
3	.50+	.30+	.25	1.05

TABLE XI

HYPOTHETICAL PROBABILITY ESTIMATES FOR CLASSES GIVEN SINGLE CUES

Class	Cue A	Cue B	Cue C		
Bayesian Model					
1	(.80)	(.01)	(.80)	=	.006
2	(.19)	(.65)	(.20)	=	.025
3	(.01)	(.30)	(.01)	=	.00003
Additive Model					
1	.80+	.01+	.80	=	1.61
2	.19+	.65+	.20	=	1.04
3	.01+	.30+	.01	=	.32

of the specific Bayesian equation to the experimental events cannot be made. That is, the subjects in the non-Bayesian analysis group who reported unequal class proportions did not have to incorporate their beliefs and could have obtained high consistency correlations by simply averaging.

The five Bayesian subjects in experiment (1) who obtained significant consistency correlations were among the six subjects from the Bayesian analysis group who also reached significance using the additive model. It is therefore impossible to know what process these subjects were using to combine their judgments of single cues in order to make combined-cue estimates. The examination of the Bayesian hypothesis is of course not influenced by these findings.

In summary, the present study did not support the Bayesian hypothesis. Of interest, however, was the discovery that the experimental approach utilizing Bayes's Theorem as a proportion and evaluating the postulated Bayesian revision consistency relationship by correlational analysis may not serve to distinguish the model being examined from a simple averaging process. Support of the Bayesian hypothesis based on subjects' inaccurate subjective probability estimates could be of little value, depending on the estimates that were made. It seems likely that positive results might also be explained in terms of an averaging process.

Future investigations, using the same experimental approach, should attempt to structure the subjective estimates of the subjects. This could be accomplished by using objective probabilities which would distinguish between the Bayesian and additive models and allowing subjects to develop accurate notions about the probabilistic rules. Beach (1966) states that accurate subjective probabilities would constrain subjects to be consistent. It seems clear, at this point, that the model predicting the resulting consistency could depend upon the objective probability values. That is, certain numerical values would result in high consistency using either model and other values would serve to distinguish between them.

CHAPTER V

SUMMARY

The purpose of the study was to examine the use of Bayes's Theorem of mathematical probability theory as the appropriate normative model for the manner in which subjects revise their opinions. The subjects' opinions consisted of concepts formed about the probabilistic relationships between experimental cues and classes. The cues, presented on 5 x 8 index cards, consisted of the physical characteristics, i.e., color, body shape, and leg shape, which made up a figure called a spaceman. The classes were the planets, i.e., Mars, Pluto, and Moon, to which the spacemen were assigned. The physical characteristics of the spacemen were probabilistic indicants of the planets to which they belonged.

The 57 college undergraduates who served as subjects were given only sufficient training with the experimental universe to form inaccurate notions about the probabilistic rules which linked cues to classes. The training consisted of showing the subjects the cards possessing one of each cue and thereby a whole spaceman, having them make class judgments and then identifying the correct class. After training, the subjects' subjective probability estimates were obtained for the class membership of test stimuli that possessed single cues and test stimuli that possessed a combination of three of these same cues.

The specific hypothesis was: subjective probability estimates for combined-cue stimuli are highly consistent Bayesian combinations of the subjective probability estimates for the corresponding single-cue stimuli. Highly consistent was defined in terms of being relative in the manner described by Bayes's Theorem. To evaluate consistency for each subject, three single-cue estimates were substituted into Bayes's Theorem and the resulting value correlated with the subjects' judgement for all three cues seen together.

The basic Bayesian equation was reduced to a proportion in accordance with a previous study supporting the hypothesis. The equation used is based on the subjects' believing that the classes were equally represented and therefore equally probable. Two experimental groups were used in the present study; one group was asked after the experiment to report their notions about the probability of the classes and the other group was told before that they would see an equal number of spacemen from each planet.

The results of the study did not support the Bayesian hypothesis. In addition, it was discovered that the mathematical basis upon which the hypothesis was examined cannot, with any certainty, distinguish between the model being tested and a simple averaging process.

REFERENCES

REFERENCES

Allport, F. H. *Theories of perception and the concept of structure.* New York: Wiley, 1955.

Beach, L. R. Cue probabilism and inference behavior. *Psychological Monograph,* 1964, 78, 5, (Whole No. 582).

Beach, L. R. Accuracy and consistency in the revision of subjective probabilities, *I.E.E.E. Transactions in Human Factors in Electronics,* 1966, HFE-7, 29-37.

Brunswik, E. Organismic achievement and environmental probability, *Psychological Review,* 1943, 50, 255-272.

Brunswik, E. *Systematic and representative design of psychological experiments with results in physical and social perception.* Berkeley: University of California Press, 1947.

Brunswik, E. The conceptual framework of psychology, *International Encyclopedia of Unified Science.* Chicago: University of Chicago Press, 1, No. 10, 1952.

Brunswik, E. *Perception and the representative design of Experiments.* Berkeley: University of California Press, 1956.

Edwards, W. (Guest Editor) Revision of opinions by men and man-machine systems. *I.E.E.E. Transactions in Human Factors in Electrons,* 1966, HFE-7, 29-37.

Edwards, W., Lindman, L., & Phillips, L. D. Emerging technologies for making decisions. In *New Directions in Psychology* II, New York: Holt, Rinehart, and Winston, Inc., 1965.

Goodnow, R. E. Utilization of partially valid cues in perceptual identification. Unpublished doctoral dissertation, Harvard University, 1954. Cited in Bruner, J. S., Goodnow, J. J., and Austin, G. A. *A study of thinking.* New York: Wiley, 1956.

Summer, S. A. Learning of responses to multiple weighted cues, *Journal of Experimental Psychology,* 1962, 64, 29-34.

Stilson, D. W. Probability and statistics in psychological research and theory. San Francisco: Holden-Day, Inc., 1966.

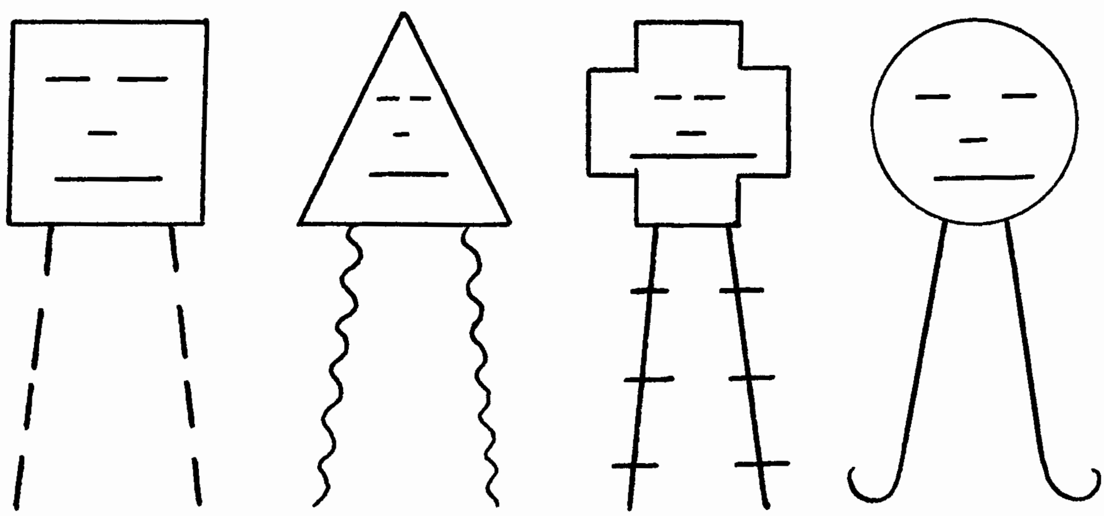
Uhl, C. N. Learning interval concepts: I. Effects of differences in stimulus weights. Journal of Experimental Psychology, 1963, 66, 264-273.

APPENDIX I

APPENDIX I

EXPERIMENTAL UNIVERSE

The cues formed objects referred to as spacemen. The spacemen were divided into the classes of Mars, Pluto and Moon. The cues were the body shapes of circle, square, triangle and cross; the leg shapes of dashed lines, wavy lines, straight lines with cross marks, and continuous lines which slanted out and curved up at the bottom; and the body colors of red, yellow, blue and green. The figures below show all of the cues except those in the color dimension.



APPENDIX II

APPENDIX II

THE CUE COMBINATIONS AND CLASS ASSIGNMENTS THAT COMPRISED THE EXPERIMENTAL UNIVERSE*

Cue Combinations	Class	Cue Combinations	Class
A1 I	Pluto	C1 I	Mars
A1 II	Mars	C1 II	Mars
A1 III	Moon	C1 III	Moon
A1 IV	Moon	C1 IV	Mars
A2 I	Pluto	C2 I	Moon
A2 II	Moon	C2 II	Mars
A2 III	Moon	C2 III	Moon
A2 IV	Moon	C2 IV	Moon
A3 I	Pluto	C3 I	Pluto
A3 II	Mars	C3 II	Mars
A3 III	Pluto	C3 III	Moon
A3 IV	Pluto	C3 IV	Moon
A4 I	Pluto	C4 I	Mars
A4 II	Mars	C4 II	Mars
A4 III	Pluto	C4 III	Moon
A4 IV	Pluto	C4 IV	Moon
B1 I	Pluto	D1 I	Mars
B1 II	Mars	D1 II	Mars
B1 III	Moon	D1 III	Mars
B1 IV	Pluto	D1 IV	Mars
B2 I	Pluto	D2 I	Pluto
B2 II	Moon	D2 II	Mars
B2 III	Moon	D2 III	Moon
B2 IV	Moon	D2 IV	Moon
B3 I	Pluto	D3 I	Mars
B3 II	Pluto	D3 II	Mars
B3 III	Pluto	D3 III	Moon
B3 IV	Pluto	D3 IV	Pluto
B4 I	Pluto	D4 I	Mars
B4 II	Mars	D4 II	Mars
B4 III	Pluto	D4 III	Moon
B4 IV	Pluto		

*Cue dimension key

Body shapes:		Body colors:		Leg shapes:	
Circle	- A	Red	- 1	Dashed	- I
Cross	- B	Yellow	- 2	Wavy	- II
Triangle	- C	Blue	- 3	Cross marks	- III
Square	- D	Green	- 4	Curved up	- IV

APPENDIX III

APPENDIX III

TEST TRIAL CUES AND CUE COMBINATIONS

Test card no.	Cues and cue combinations
1*	Body shape--square
2*	Color--green
3	Leg shape--wavy lines
4	Body shape--circle, color--green, leg shape--dashed lines
5	Leg shape--curved up lines
6*	Body shape--triangle
7*	Color--yellow
8	Body shape--triangle, color--red, leg shape--wavy lines
9	Body shape--circle
10*	Leg shape--dashed lines
11	Color--red
12	Body shape--square, color--yellow, leg shape--curved up lines

*Presented to the subjects a second time for a reliability measurement.

APPENDIX IV

APPENDIX IV

INDIVIDUAL SUBJECT'S RESPONSE FORM

Name______________________________

Date______________________________

Group_____________________________

1.	MARS PLUTO MOON	10.	MARS PLUTO MOON	19.	MARS PLUTO MOON	28.	MARS PLUTO MOON	37.	MARS PLUTO MOON
2.	MARS PLUTO MOON	11.	MARS PLUTO MOON	20.	MARS PLUTO MOON	29.	MARS PLUTO MOON	38.	MARS PLUTO MOON
3.	MARS PLUTO MOON	12.	MARS PLUTO MOON	21.	MARS PLUTO MOON	30.	MARS PLUTO MOON	39.	MARS PLUTO MOON
4.	MARS PLUTO MOON	13.	MARS PLUTO MOON	22.	MARS PLUTO MOON	31.	MARS PLUTO MOON	40.	MARS PLUTO MOON
5.	MARS PLUTO MOON	14.	MARS PLUTO MOON	23.	MARS PLUTO MOON	32.	MARS PLUTO MOON	41.	MARS PLUTO MOON
6.	MARS PLUTO MOON	15.	MARS PLUTO MOON	24.	MARS PLUTO MOON	33.	MARS PLUTO MOON	42.	MARS PLUTO MOON
7.	MARS PLUTO MOON	16.	MARS PLUTO MOON	25.	MARS PLUTO MOON	34.	MARS PLUTO MOON	43.	MARS PLUTO MOON
8.	MARS PLUTO MOON	17.	MARS PLUTO MOON	26.	MARS PLUTO MOON	35.	MARS PLUTO MOON	44.	MARS PLUTO MOON
9.	MARS PLUTO MOON	18.	MARS PLUTO MOON	27.	MARS PLUTO MOON	36.	MARS PLUTO MOON	45.	MARS PLUTO MOON

46. MARS
PLUTO
MOON

47. MARS
PLUTO
MOON

48. MARS
PLUTO
MOON

49. MARS
PLUTO
MOON

50. MARS
PLUTO
MOON

51. MARS
PLUTO
MOON

52. MARS
PLUTO
MOON

53. MARS
PLUTO
MOON

54. MARS
PLUTO
MOON

55. MARS
PLUTO
MOON

56. MARS
PLUTO
MOON

57. MARS
PLUTO
MOON

58. MARS
PLUTO
MOON

59. MARS
PLUTO
MOON

60. MARS
PLUTO
MOON

61. MARS
PLUTO
MOON

62. MARS
PLUTO
MOON

63. MARS
PLUTO
MOON

64. MARS
PLUTO
MOON

65. MARS
PLUTO
MOON

66. MARS
PLUTO
MOON

67. MARS
PLUTO
MOON

68. MARS
PLUTO
MOON

69. MARS
PLUTO
MOON

70. MARS
PLUTO
MOON

71. MARS
PLUTO
MOON

72. MARS
PLUTO
MOON

73. MARS
PLUTO
MOON

74. MARS
PLUTO
MOON

75. MARS
PLUTO
MOON

76. MARS
PLUTO
MOON

77. MARS
PLUTO
MOON

78. MARS
PLUTO
MOON

79. MARS
PLUTO
MOON

80. MARS
PLUTO
MOON

81. MARS
PLUTO
MOON

82. MARS
PLUTO
MOON

83. MARS
PLUTO
MOON

84. MARS
PLUTO
MOON

85. MARS
PLUTO
MOON

86. MARS
PLUTO
MOON

87. MARS
PLUTO
MOON

88. MARS
PLUTO
MOON

89. MARS
PLUTO
MOON

90. MARS
PLUTO
MOON

91. MARS
PLUTO
MOON

92. MARS
PLUTO
MOON

93. MARS
PLUTO
MOON

94. MARS
PLUTO
MOON

95. MARS
PLUTO
MOON

96. MARS
PLUTO
MOON

97. MARS
PLUTO
MOON

98. MARS
PLUTO
MOON

99. MARS
PLUTO
MOON

100. MARS
PLUTO
MOON

101. MARS
PLUTO
MOON

102. MARS
PLUTO
MOON

103. MARS
PLUTO
MOON

104. MARS
PLUTO
MOON

105. MARS
PLUTO
MOON

106. MARS
PLUTO
MOON

107. MARS
PLUTO
MOON

108. MARS
PLUTO
MOON

109. MARS
PLUTO
MOON

110. MARS
PLUTO
MOON

111. MARS
PLUTO
MOON

112. MARS
PLUTO
MOON

113. MARS
PLUTO
MOON

114. MARS
PLUTO
MOON

115. MARS
PLUTO
MOON

116. MARS
PLUTO
MOON

117. MARS
PLUTO
MOON

118. MARS
PLUTO
MOON

119. MARS
PLUTO
MOON

120. MARS
PLUTO
MOON

121. MARS
PLUTO
MOON

122. MARS
PLUTO
MOON

123. MARS
PLUTO
MOON

124. MARS
PLUTO
MOON

125. MARS
PLUTO
MOON

126. MARS
PLUTO
MOON

SAMPLE TRIAL

		No Probability		Certainty
	MARS	0	________	100
		No Probability		Certainty
	PLUTO	0	________	100
		No Probability		Certainty
	MOON	0	________	100

- -

		No Probability		Certainty
1.	MARS	0	________	100
	PLUTO	0	________	100
	MOON	0	________	100
		No Probability		Certainty
2.	MARS	0	________	100
	PLUTO	0	________	100
	MOON	0	________	100
		No Probability		Certainty
3.	MARS	0	________	100
	PLUTO	0	________	100
	MOON	0	________	100
		No Probability		Certainty
4.	MARS	0	________	100
	PLUTO	0	________	100
	MOON	0	________	100

		No Probability		Certainty
5.	MARS	0	______	100
	PLUTO	0	______	100
	MOON	0	______	100

		No Probability		Certainty
6.	MARS	0	______	100
	PLUTO	0	______	100

		No Probability		Certainty
7.	MARS	0	______	100
	PLUTO	0	______	100
	MOON	0	______	100

		No Probability		Certainty
8.	MARS	0	______	100
	PLUTO	0	______	100
	MOON	0	______	100

		No Probability		Certainty
9.	MARS	0	______	100
	PLUTO	0	______	100
	MOON	0	______	100

		No Probability		Certainty
10.	MARS	0	______	100
	PLUTO	0	______	100
	MOON	0	______	100

		No Probability		Certainty
11.	MARS	0	______	100
	PLUTO	0	______	100
	MOON	0	______	100

No Probability — Certainty

12. MARS 0 ________________ 100

PLUTO 0 ________________ 100

MOON 0 ________________ 100

No Probability — Certainty

13. MARS 0 ________________ 100

PLUTO 0 ________________ 100

MOON 0 ________________ 100

No Probability — Certainty

14. MARS 0 ________________ 100

PLUTO 0 ________________ 100

MOON 0 ________________ 100

No Probability — Certainty

15. MARS 0 ________________ 100

PLUTO 0 ________________ 100

MOON 0 ________________ 100

No Probability — Certainty

16. MARS 0 ________________ 100

PLUTO 0 ________________ 100

MOON 0 ________________ 100

No Probability — Certainty

17. MARS 0 ________________ 100

PLUTO 0 ________________ 100

MOON 0 ________________ 100

APPENDIX V

APPENDIX V

INSTRUCTIONS FOR TRAINING TRIALS

Most college men who are over 6 feet by 8 inches tall and who wear letterman sweaters are basketball players; but some are not. If we were forced to make a decision based on the information given, we would probably classify such an individual as a basketball player. Our daily lives are filled with choices based on uncertain information. For example, does tail wagging indicate that this dog is friendly? Does an east wind and dark skies indicate rain?

This experiment will investigate how we make these kinds of judgements. It will involve a kind of game in that we are going to pretend that three spaceships have landed on campus. The spaceships are each from a different planet; one from Mars, one from Pluto, and one from the Moon. You will be shown cards containing pictures of the spacemen. Like this . . . . Each card contains some of the information you will need to tell where the particular spaceman came from, i. e., Mars, Pluto or Moon. Your task in the experiment is to learn how to use the information just as you have learned that height and letterman sweaters are good indicators of basketball players.

You will see each card for five seconds. After I remove it from view you must immediately classify the spaceman according to his planet by marking the preprinted answer sheet. There will be a three second delay before I identify the correct planet so that you may be sure to make a choice. At the beginning you will have little upon which to base a choice; you will be only guessing; make a choice regardless. Even if you are not at all sure, answer each time.

This is how it will work. Please do not mark your answer sheet. First I will show you the card for five seconds--do not answer during this time; then I will remove it from view for three seconds and say, "Circle your answer." You must mark an answer during this time. Then I will show the card to you again and identify the correct planet. Please do not mark your paper or change your choice after the correct planet is identified. It is important that you realize that while your task is to learn how to use the information provided, the information is of an uncertain nature. Just as with the cues that helped you identify the basketball player, they may be correct most of the time, but may also be wrong on occasion.

APPENDIX VI

APPENDIX VI

INSTRUCTIONS FOR TEST TRIALS

In this part of the experiment I am going to show you 12 cards. On some of the cards you will not be able to see all of the spaceman, i.e., you will only see certain parts of him. You will see only his body shape with no color; you will see only his leg shape and you will see his color without being able to determine his body and leg shapes. On other cards you will see the whole spaceman.

When you are shown the information on the card, whether it is part of, or the complete spaceman, you will be asked to estimate the probability that the information is an indicator for each class. In other words, if you see the color blue, for example, you will estimate the probability that "blue" indicates a spaceman is from Mars, then from Pluto, then from Moon.

Don't let the word probability bother you. Even if you don't like mathematics and haven't had a course since you were a freshman in high school, you can perform these tasks.

I will define the word probability for our purposes to be sure that we are all agreed on what it means. We will use probability to mean proportion or percentage. To begin with, consider a deck of 52 playing cards. We know that the cards in the deck are of several different kinds. Suppose that we are interested in the spade as a kind of card. The probability of a spade in the deck is the number of times this card occurs in the deck divided by the number of cards in the deck. In other words, the proportion or percentage of cards that are spades. There are thirteen spades in a deck of 52 cards. Thirteen is one fourth of 52, so the probability of getting a spade is 1/4, .25 or 25 per cent. Now consider a bag of 100 marbles--50 black and 50 white. What is the probability of getting a black marble? The black marbles consist of one half of all the marbles; therefore the probability of getting a black one is 1/2, .50 or 50 per cent. What is the probability of getting a green marble from the same sack? It is of course zero, or no probability, because there are no green marbles. In a sack of 100 black marbles what is the probability of drawing a black marble? The answer is, of course, 1.00, 100 per cent or certainty because all of the marbles are black. If any of you are having difficulty with the meaning of probability you can think of it directly in terms of percentages. For example, you may expect to draw a white marble from a sack of half white and half black 50 per cent of the time. Are there any questions?

When you are shown the information on the card you will make your estimate of probability by making a mark on the straight line representing the full range of probability. Next you will place a number by your mark showing the number value of your estimate. Please notice that the straight line has a "0" on the left and a "100" on the right. Probabilities and percentages are usually written with decimals, as they were in the examples I have given, with 1.00 being certainty. We are going to use 100 as the perfect probability or certainty just to keep from using decimals. In other words, instead of writing .10 you will write 10; instead of writing .90, you will write 90, and so on. If, for example, you were shown the body shape of a cross and you estimated the probability of this shape being an indicator of Mars at around one third, you would place a mark approximately one third of the way along the line and place the number 33 at the mark. Are there any questions?

I will show you a card and ask you to estimate the probability that the information you see is an indicator for each of the three classes. The card will remain in view during all three estimates. It will work like this. . . . I will ask you to make an estimate for Mars, Pluto and Moon. You will have twenty seconds and then we will go on to the next card. Here is a sample card to be sure you understand. Please mark your answers under the sample trial on your answer sheet. Each of you should have a mark and a number for each planet. Are there any questions?

*1. Estimate the probability that the body shape square indicates a spaceman is from Mars, Pluto, Moon.

*2. Estimate the probability that the color green indicates a spaceman is from Mars, Pluto, Moon.

3. Estimate the probability that this leg shape indicates a spaceman is from Mars, Pluto, Moon.

4. Estimate the probability that this spaceman is from: Mars, Pluto, Moon.

5. Estimate the probability that this leg shape indicates a spaceman is from Mars, Pluto, Moon.

*6. Estimate the probability that the body shape triangle indicates a spaceman is from Mars, Pluto, Moon.

*7. Estimate the probability the color yellow indicates a spaceman is from Mars, Pluto, Moon.

8. Estimate the probability that this spaceman is from Mars, Pluto, Moon.

9. Estimate the probability that the body shape circle indicates a spaceman is from Mars, Pluto, Moon.

*10. Estimate the probability that this leg shape indicates a spaceman is from Mars, Pluto, Moon.

11. Estimate the probability that the color red indicates a spaceman is from Mars, Pluto, Moon.

12. Estimate the probability that this spaceman is from Mars, Pluto, Moon.

*Ask the Ss to make estimates for test cards 1, 2, 6, 7, and 10 again.

APPENDIX VII

APPENDIX VII

SUBJECT QUESTIONNAIRE

NAME (important) ______________________________

In the experiment a total of 126 spacemen were identified according to the planets MARS, PLUTO and MOON. Try to recall your impression of the approximate number from each planet, i.e., did you think there were more from one planet than another or did you get the idea that there was an equal number from each planet? (Circle your answer.)

A. There were more spacemen from one planet than another, i.e., an unequal breakdown.

B. There was approximately an equal number of spacemen from each planet.

C. Can't remember.

D. Did not have an idea at the time.

COMMENTS: