1966

A Comparison of Estimates of Relative Frequency vs. Subjective Probability

Duncan Muir McQuarrie
Central Washington University

15

A COMPARISON OF ESTIMATES OF RELATIVE FREQUENCY VS. SUBJECTIVE PROBABILITY

A Thesis

Presented to

the Graduate Faculty

Central Washington State College

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

Duncan Muir McQuarrie

August, 1966

APPROVED FOR THE GRADUATE FACULTY

____
Jack J. Crawford, COMMITTEE CHAIRMAN

____
Thomas B. Collins Jr.

____
Maurice L. Pettit

A COMPARISON OF ESTIMATES OF RELATIVE FREQUENCY VS. SUBJECTIVE PROBABILITY

The purpose of this study is to extend the investigation of the general area of subjective probability by exploring three methods of asking for estimates. Inasmuch as this area has not been a major focus of interest for experimental studies, a brief review of the kinds of studies emanating from subjective probability notions will be given.

Edwards (1954) distinguishes between two uses of the phrase "subjective probability." The first, a school of thought about the logical basis of mathematical probability and the second, a name for a transformation on the scale of mathematical probabilities which is somehow related to behavior. De Finetti (1937) is considered to have been the first to give serious consideration to a formalized theory of subjective probability. This publication is substantially devoted to a development of subjective probability as the basis for, or logical foundation of mathematical probability. Savage (1954) follows De Finetti and is also devoted to the development of subjective probability as the basis for mathematical probability. Both are formalized approaches to the theoretical aspects of subjective probability and deal little with empirical studies.

Subjective probability in the second usage refers to an opinion in the form of a measure of personal belief in a particular proposition. The proposition concerns the occurrence of a certain event, with the degree of belief usually expressed as a number or ratio. It is this second usage that is the interest of this paper. While subjective probability has been a key concept in studies of decision theory, little has been done to investigate the behavioral implications of such subjective probability concepts. Where investigation has been attempted, the methods of obtaining estimates have been varied and usually indirect.

Luce, Bush, and Galanter (1965) attribute the first attempt to measure subjective probability experimentally to Preston and Baratta (1948). In this study subjects used play money to bid for a card on which a probability was written. The highest bidder was then allowed a chance to roll a set of dice with the probability of winning corresponding exactly to the probability printed on the card presented for auction. Subjective probability was computed indirectly as demonstrated by the following example. On a given play, subjects may have bid for a prize of 250 points with probability 0.25 of winning. If the average successful bid was 50, then subjective probability was computed to be 50/250 = 0.20. Using this method of computation it was concluded that subjects tend to underestimate when the objective probability is greater than 0.20 and systematically over-estimate when the objective probability is less than 0.20.

This method has a number of faults. While it is not clear whether or not there were restrictions on the bids, it is quite possible that restrictions could have existed in a variety of forms. It has been hypothesized (Edwards, 1953) that subjects tend to have preferences for certain probabilities over others. If this is a valid relationship, then it could have suppressed subject's bidding behavior for other than preferred probabilities. In other words, this could have resulted in self-imposed restrictions on the subject's bids. Another possible restriction on the bids was the total amount available at any given time for use in bidding. It would seem quite likely that subjects with small total amounts with which to bid from would be more hesitant to bid than subjects with an abundance with which to bid from. Because winning the bid does not necessarily mean winning the bet, subjects could have depleted their total amount to the point that they may have entered into a bidding situation with inadequate amounts with which to bid. These types of restrictions on a bid would result in a marked confounding of a computed subjective probability.

Edwards (1961) gives the two approaches most commonly used to measure subjective probability as inference from a Subjective Expected Utility (SEU) model, the SEU model being a model for decision theory, and direct psychophysical judgment methods. In the SEU model subjective probability is usually inferred but not without being confounded with the concept of utility; utility being the value the occurrence of the event has for the individual. Davidson, Suppes, and Siegal (1957), in their book on decision theory, devote a chapter to a discussion of this method of measuring subjective probability.

Subjects were allowed to choose between two courses of action (see Figure 1). The two states controlled which

	state 1	state 2
option 1	win 15¢	lose 8¢
option 2	lose 8¢	win 15¢

Figure 1. Matrix used by Davidson, et al. (1957)

pay off was given and were determined by some random device. Various random devices were tried until one was found so that the subjects were indifferent about the two options. If they were indifferent to the options, then this was felt to imply that the subjective probabilities of the two events were equal; i.e., equal to 0.50. Having found events with equal subjective probabilities (a die on three faces of which they had printed the nonsense syllable ZOJ and the syllable ZEJ on the other three faces), they proceeded to use these in finding a function representing utility values. After finding this function, they then used it and the same bets to find the subjective probability of other events.

Using the utility function to find the subjective probability of other events appears to have several serious limitations. First, in constructing the equally-spaced utility intervals using the subjective probability of 0.50, it is assumed that this subjective probability remains at 0.50 regardless of what outcomes are paired with it and regardless of experience with it (Edwards, 1961). Anything that might have caused the subjective probability to change from 0.50 would invalidate the constructed utility function which, in turn, would have made a derived subjective probability based on it invalid. A second limitation is the possibility of an interaction between utility and subjective probability. Edwards (1961) points to this relationship in interpreting the results of measuring subjective probability in a utility measurement experiment.

> Subjective probability functions obtained from bets on which subjects would only win or break even indicate that subjective probability exceeded objective probability at all points between 0 and 1. But functions obtained from bets on which subjects could only lose or break even indicate that the subjective probability equalled objective probability. In other words, there was a vigorous interaction between the sign of the pay off and the shape of the subjective probability function. (Edwards, 1961, p. 480.)

Irwin (1953) suggests that the degree of subjective probability and utility, not merely the sign or pay off, interact. If this is the case, then any attempt to measure subjective probability in the context of the SEU model is going to be confounded with the interaction between utility or sign and the subjective probability.

Luce, et al. (1965) describe another game-like method used by Shuford (1959) to obtain measurements of subjective probability. This technique was like the one used in the Preston and Baratta (1948) study in that it was a bidding situation. It differed to the extent that the subjects did not bid against each other and the bids were based upon a 20 x 20 matrix of vertical and horizontal bars. The subjects were run in pairs with one subject bidding that a horizontal or vertical bar would be selected while the other subject decided to buy or sell that bid. The bid was kept between 0 and 10. The experimenter would randomly select two intergers from the intergers 1 to 20 in order to pick a row and column of the matrix to enter. Shuford (1959) theorized that the bidders best strategy in the game was to set the bid equal to ten times his subjective probability of the favorable outcome. Shuford transformed the data by dividing the bid by 10. In so doing Shuford claimed to obtain a number representing the subject's subjective probability. Such results from Shuford's subjects confirmed the earlier findings of Preston and Baratta in that a large fraction of the subjects over-estimated low probabilities and underestimated high ones. Shuford also found that subjective probability estimates of a number of the subjects were fit quite well by a linear function of objective probability.

This method has some of the limitations of the Preston and Baratta (1948) study. Direct limitations are placed on the subject's bid; that is, the bid must be from 1 to 10. When the upper limit of the bid is 10, the computed subjective probability (10/10) could never be greater than 1, a relationship that has been suggested to not necessarily exist (Edwards, 1961).

It is not clear whether the subjects had an unlimited amount with which to bid or if this was an initial amount which decreased or increased as a result of "playing the game." It is possible that bidding and therefore computed subjective probability interacts with the total amount available at the time of the bid. Another source of confounding was present in the form of inter-personal interactions. While the bidding situation was not a competitive one, such as in the Preston and Baratta study, the subjects were run in pairs with one subject bidding and the other subject buying or selling that bid. It is quite possible that the choosing to buy or sell another's bid functioned as a reinforcement for that other person. This form of social reinforcement in the interaction situation might have influenced the size of succeeding bids. The transformation of the data by Shuford would also seem to be open to criticism. No explanation is given as to why the bids were divided by the specific value of 10. An obvious explanation would be that this constant produced the best approximation of the objective probability, a technique that would be highly suspicious.

Each of the foregoing methods of measuring subjective probability has been related to a game or uncertain outcome situation. In all of them the measured or computed subjective probability has been confounded with; limits of bids, interactions with utility, inter-personal interactions, transformations of the data, or some combination of these.

An alternative approach is that of direct psychophysical measurement. Edwards (1961), in referring to direct psychophysical judgment methods, describes them as requiring subjects to estimate the proportion of one class of elements in a display that has more than one type of element. While he describes this technique as the "estimating of proportion," he refers to it as a "direct" method of measuring subjective probability as opposed to the "indirect" method of inference from an SEU model. Luce, et al. (1965), in referring to the same general class of experiments, that is, those referred to as psychophysical judgment methods, suggest that while they are not explicitly concerned with the measurement of subjective probability, they do bear a rather close relation to it. This relation is seen as the scale of proportion that is based on the responses of subjects to randomly composed arrays of two or more types of elements. This suggests that subjective probability is based upon the same type of information as mathematical probability. That is, while mathematical probability is a function of relative frequencies or the ratio of the number of units in each class to the number of units in the whole population, subjective probability is based upon "estimates" of the relative frequencies. These direct psychophysical judgment methods usually require subjects to estimate the proportion of one type of stimulus element of two types.

Stevens and Galanter (1957), while investigating ratio scales and category scales, had subjects estimate percentage of blue and green dots. Subjects were shown cards each containing 36 dots in the two colors. The number of dots of one color varied from 3 to 33 in steps of three. The subjects were shown the card and then asked to estimate percentage of blue or green dots. Each subject was asked to give estimates for each color. The results of this study showed the relation between average responses and proportion was an inverse ogive. Stevens and Galanter (1957) suggest these results indicate subjects are fairly accurate estimators of proportion.

Shuford (1961) felt that subjects in the Stevens and Galanter (1957) study may have been able to count the dots of the smaller proportion when they number 3 to 6, thus producing an inverted ogive curve. When making percentage estimates, Shuford felt that subjects could figure that 3/36 = 1/12 is about 8 per cent and that 33/36 = 1-1/12 is about 92 per cent. This would give almost no error at the extremes; however, there would be a certain amount of uncertainty at the middle as subjects would have a difficult time counting the larger numbers. This type of relation would produce the inverse ogive. Using the same method of presenting elements as Shuford (1959), the number of elements presented was increased to 400. When using a small proportion of 10 per cent, the relative frequency was 40 and Shuford (1961) felt that this relative frequency made it impossible to count the elements and arrive at an estimate of proportion. The results showed the relation between percentage estimate and proportion to be an S-shaped ogive rotated slightly about the center of the graph in a clockwise direction. This is different from the results of Stevens and Galanter (1957) who obtained results yielding an _inverse_ ogive.

Holmberg (1964), in an investigation of the additivity theorem as applied to subjective probability used a direct psychophysical technique of obtaining judgments of proportion. Subjects viewed a series of slides showing various complex objects, such as: airplanes, birds, flowers, and ceramics. Following the presentation of the slides, subjects were given a questionnaire containing items related to several aspects of the slides. On each questionnaire only one of the questions was related to the task of estimating proportion as a measure of subjective probability. It was assumed that by having the question on proportion presented with other non-related questions that the task of the experiment remained concealed and that subjects did not count objects and arrive at estimates on succeeding trials. The results of this study were non-significant due to large amounts of individual variance. While concealing the task from the subjects might have helped prevent them from counting, it would also seem to have encouraged subjects to guess what the task was going to be. Subjects attempting to anticipate the experimenter might have concentrated on inappropriate stimulus dimensions producing inaccurate estimates not because of inability to judge proportion but because of interference from concentrating on characteristics other than observed relative frequency.

A number of techniques of measuring subjective probability have been described. Each of these approaches has had some limitations in terms of measuring subjective probability. In the early studies using game models (Preston and Baratta, 1948), subjective probability was inferred from a betting or bidding situation in which subjective probability was compounded with restrictions on the bid and availability of money to bid with. In the SEU model (Davidson, Suppes, Siegal, 1957) measurement of subjective probability was compounded with utility. In the technique used by Shuford (1959) a subjective probability derived from subjects' bids was confounded with social reinforcement as well as the apparent arbitrary transformation of the data so as to produce results approximating the objective proportions. While direct psychophysical measurement would seem to eliminate some of the compounding involved in game or uncertain outcome models, there have been limitations with the methods used. Direct psychophysical judgments of observer proportion is a task claimed to be related to estimates of subjective probability. Stevens and Galanter (1957) had subjects estimate proportion of dots contained in a stimulus field. Shuford (1961) felt that the relative frequencies of the smaller proportions in the Stevens and Galanter study allowed subjects to count the elements and thus arrive at the proportion mathematically. By increasing the number of elements to 400 and having subjects estimate proportion, Shuford obtained results somewhat different from Stevens and Galanter. Subjects in both these studies were asked for estimates of proportion reported as a percentage. Holmberg (1964) may have confounded his results by concealing the task of estimating proportions.

While the task of asking for estimates of proportion seems to have some promise for investigating subjective probability, the relation between observed proportion and subjective probability needs to be further understood. Shuford (1961) points to the necessity of understanding the task of estimating proportions where the same type stimuli are to be used for estimating subjective probability. In attempting this understanding, Shuford (1961) investigated several parameters of asking for percentage estimates. Among those studied were element type and length of exposure time of the stimuli.

It may be that there are differences in estimates due to the way we ask for those estimates. There have been no reports of investigations comparing asking for estimates of subjective probability with asking of estimates of observed proportion when those estimates were based on the identical set of stimuli.

HYPOTHESES

If, as Edwards (1961) suggests, "subjective probability is linearly related to observed proportion," then estimates of relative frequency should be equal to a simple linear function of estimates of subjective probability. Inasmuch as this relationship has not been tested, this study was designed to test the following hypotheses:

If estimates of observed proportion, estimates of proportion expected in the future, and estimates of subjective probability are based upon identical stimuli, then there will be no difference in those estimates. Subjective probability is defined as an estimate of the likelihood of occurrence of a particular event. The specific null hypotheses tested were:

1. Estimates of subjective probability will not differ from estimates of observed proportion.

2. Estimates of subjective probability will not differ from estimates of expected long-run proportion.

3. Estimates of observed proportion will not differ from estimates of expected long-run proportion.

METHOD

Ninety subjects (Ss) were given six (6) trials. One trial consisted of viewing a set of five (5) slides, each slide showed a different 72-unit matrix of three different geometric figures. Following each trial, Ss were required to give estimates either of observed proportion, expected long-run proportion, or subjective probability. On three trials, the estimates were based on geometric figures whose mathematical proportion was .67 and on three trials estimates were based on geometric figures whose mathematical proportion was .22.

Design

A 3 x 2 randomized group design with repeated measures was used. The factors were as follows:

A. Type of estimate asked for; A_1 observed proportion; A_2 probability of occurrence of single geometric figure; A_3 expected long-run proportion.

B. Order of geometric figures or objective proportion which estimates were based upon; B_1--square (.22)--triangle (.67)--triangle (.67) square (.22)--triangle

(.67)--square (.22), B_2--triangle (.67)--square (.22)--square (.22)--triangle (.67)--square (.22)--triangle (.67).

C. Trials (repeated measures). (See Table 1.)

Subjects

Subjects were 90 undergraduate volunteers attending Central Washington State College and were randomly assigned to treatment groups.

Apparatus and Materials

Six (6) identical sets of slides containing five (5) 35 mm 2 x 2 slides were used. Each of the five slides showed seventy-two (72) geometric figures (triangles, squares, and circles of different proportions) arranged in a different non-rectangular matrix (see Appendix A). Different non-rectangular matrices were used in an attempt to prevent Ss from determining the total number of geometric figures and thus being able to use this as a basis for their estimates. The relative frequency and objective proportions for each matrix were as follows: circles - 8 - .11, squares - 16 - .22, and triangles - 48 - .67. The matrices were drawn on

TABLE I

3 x 2 RANDOMIZED GROUP DESIGN WITH REPEATED MEASURES

		TRIALS C					
		C_1	C_2	C_3	C_4	C_5	C_6
OBSERVED PROPORTION A_1	B_1	.22	.67	.67	.22	.67	.22
	B_2	.67	.22	.22	.67	.22	.67
SUBJECTIVE PROBABILITY A_2	B_1	.22	.67	.67	.22	.67	.22
	B_2	.67	.22	.22	.67	.22	.67
LONG-RUN PROPORTION A_3	B_1	.22	.67	.67	.22	.67	.22
	B_2	.67	.22	.22	.67	.22	.67

(Objective Proportions)

white paper using India ink and then photographed using a 35mm camera and "Kodak Pan-X" film. The negative prints were then made into 2 x 2 slides producing a light-on-dark projected image.

Three slides showing the three geometric figures individually were used for demonstrating the geometric figures at the beginning of the experimental session. All slides were projected upon a 5' x 5' portable screen using a "Kodak Carousel" slide projector. The slides were placed in the carousel with two blank spaces between the three separate geometric figures and the first set of five matrices, and two blank spaces between each of the six (6) sets of five (5) slides. This particular model projector, lacking a shutter, projected light upon the screen when placed on a blank space. This supplied ample lighting for the Ss to record estimates at the end of each trial. Ss viewed the slides in a 10' x 20' room which seated a maximum of nine (9) Ss at one time. To one side of the screen was a blackboard that E used while giving instructions. Each S was provided with pencil, file folder with attached envelope, and six answer sheets. The answer sheets were contained in the envelopes and Ss were required to place a completed answer sheet in the file folder at the end of each trial. Each answer sheet consisted of a slip of paper approximately 2" x 8" with the denominator of 100 shown (/100).

Procedure

The study was run on two successive days with Ss run in groups that varied in number from 3 to 9. Ss were held in a waiting room prior to the actual running of each group. Each group was randomly assigned a treatment with the exception that the final Ss were assigned treatments so as to equate the N of each group. E would enter the waiting room and lead Ss to the experimental room where they were requested to pick up the file folders and pencils from their seats before being seated. As Ss entered the experimental room, they were instructed to refrain from talking until the end of the experimental session.

After Ss were seated, E remained at the rear of the room beside the slide projector and read the following set of general instructions.

> "Please hold the file folder and pencil in your lap, do not write anything. The experiment that you are about

> to participate in is being conducted by the Department of Psychology as a part of a larger investigation of the area of decision theory. Generally, you will view some sets of slides projected upon the screen before you. Each set will contain a number of slides showing geometric figures such as these. (At this time, E projected the three separate slides showing the geometric figures.)"

In reviewing investigations of relative frequency and subjective probability, a problem common to both was evident. This problem centered on the amount of individual and within-group variance among estimates. Luce, et al. (1965) in summarizing the findings of Shuford (1959) report that some Ss' estimates of subjective probability were well fit by a linear function of objective probability, although the slope and intercept of this function varied from one subject to another. Coombs and Beardslee (1954) point to this type of variance in mentioning some unpublished analysis of data for individual subjects of Edwards (1953). Coombs and Beardslee (1954) suggest that grouping of data in studies of subjective probability may mask large individual differences. Crawford (1963), in investigating subjects' ability to make theoretical long-run estimates of frequency had considerable within-group variance, although not enough to prevent attaining significant difference between groups. Holmberg (1964), in an investigation of the additivity theorem as applied to subjective probability had subjects view different classes of objects and then asked for estimates of complimentary subjective probabilities. Individual variance was large enough to prevent attaining significant differences.

It was thought by this experimenter that individual variance in estimates could be caused by Ss lack of understanding concerning the concept of probability. In an attempt to control for this suspected cause of variance while running a follow-up study to the study run by Holmberg (1964), the following procedure was used. Ss were given a short definition of probability followed by an example prior to being asked for estimates of subjective probability. There still was a large amount of individual variance among estimates and lack of understanding concerning the concept of probability still seemed a possibility. In an attempt to control for this source of variance in the present study, E gave three different sets of instructions related to the estimates asked for. Groups received the instructions on the basis of factor A, type of estimate asked for. The instructions were matched in terms of length, wording, and examples used (see Appendix B).

The instructions related to estimates asked for were given once to each group and followed the general instructions. E went to the blackboard beside the 5' x 5' screen and used it to graphically show the numerical examples in the instructions. All examples were erased at the conclusion of the instructions. E then returned to the back of the room beside the projector and read the following set of instructions: "Please observe carefully the following set of slides, but make no notations." The first set of five (5) slides was then shown. Each slide was shown for two seconds with one second between each slide. Following the first set of five (5) slides, E asked for an estimate. This procedure of asking Ss to observe the slides, showing a set of five (5) slides, and asking for an estimate was repeated five more times for each group.

The estimate asked for on each trial was determined by the treatment group. On three of these trials Ss estimates were based upon triangles whose objective proportion was .67 and on the other three trials the estimates were based upon squares whose objective proportion was .22. To control for any order effect that might have resulted from using two different objective proportions as bases for estimates two different orders of objective proportions were used (see Table I). One order was the mirror image of the other. The forms of the questions used to elicit the estimates can be found in Appendix C.

At the conclusion of the experimental session, Ss were thanked for their participation and asked to refrain from discussing the experiment until the conclusion of the series. At this time, Ss were led from the experimental room back to the waiting room.

RESULTS

Estimates for all conditions were converted to absolute error scores to make the estimates based on different objective proportions comparable. Absolute error scores are defined thus as the absolute deviation or error of an estimate from the objective proportion that the estimate was based upon. By using "absolute" error scores, some of the artifacts of averaging were eliminated. An analysis of variance as so described by Edwards (1964) was used to analyze the data. A summary of the Analysis of Variance for the absolute error scores is shown in Table 2. Three values of F are significant, treatment and order x trials are significant beyond the .01 level, while order is significant beyond the .05 level.

The treatment effect A, type of estimate asked for, yielded an F of 6.10 which for 2 and 84 degrees of freedom is significant beyond the .01 level. Treatment effect A was partitioned into its three levels and t-tests of the difference between means computed. The mean error for estimates of subjective probability (level 2 of factor A) was 19.68. This mean was significantly greater than the means of the other two levels of factor A, estimates of observed proportion (15.13) and estimates of long-run proportion (14.63). The mean difference between estimates of observed proportion and subjective probability was 4.55 giving a t of 3.074 which is significant beyond the .01 level for 358 degrees of freedom. The mean difference between estimates of subjective probability and estimates of long-run proportion was 5.05 giving a t of 3.322 which is significant beyond the .01 level for 358 degrees of freedom. The mean difference between estimates of observed

TABLE II

ANALYSIS OF VARIANCE

SOURCE OF VARIATION		SUM OF SQUARES	d.f.	MEAN SQUARE	F
A:	Estimate Type	2,780.9	2	1390.45	6.100**
B:	Order	1,185.19	1	1185.19	5.207*
AXB:	Estimate X Order	645.87	2	322.93	1.418
	Error (a)	19,118.52	84	227.6	
C:	Trials	592.35	5	118.47	0.749
AXC:	Estimate X Trials	1,946.97	10	194.70	1.232
#BXC:	Order X Trials	9,218.88	5	1843.78	11.668**
AXBXC:	Estimate X Order X Trials	1,082.79	10	108.28	0.685
	Error (b)	66,369.35	420	158.02	
	TOTAL	102,940.82	539		

*P < .01

**P < .05

#See Table 3

TABLE III

ANALYSIS OF VARIANCE FOR PROPORTION PULLED FROM ORDER X TRIALS INTERACTION

SOURCE OF VARIATION	SUM OF SQUARES	d.f.	MEAN SQUARE	F
Proportion	8,639.99	1	8,639.99	61.004**
Proportion X Trials	1,764.08	5	352.82	2.233*
Error (b)	66,369.35	420	158.02	

**P < .01

*P < .05

TABLE IV

MEANS FOR RAW SCORES AND ABSOLUTE ERROR SCORES BY OBJECTIVE PROPORTION

TYPE OF ESTIMATE	OBJECTIVE PROPORTION BASED ON	RAW SCORE MEAN	DEVIATION FROM OBJECTIVE PROPORTION	ABSOLUTE ERROR SCORE MEAN
A_1	.22	31.20	+ 9.20	11.51
	.67	53.24	-13.76	18.76
A_2	.22	31.37	+ 9.37	14.87
	.67	52.37	-14.78	24.49
A_3	.22	27.36	+ 5.39	11.07
	.67	56.36	-10.64	18.20

proportion and estimates of long-run proportion was .5 giving a t of .362 which is non-significant.

On the basis of the significant differences obtained above, the null hypothesis is rejected for estimates of subjective probability vs. estimates of observed proportion and estimates of subjective probability vs. estimates of expected long-run proportion. The null hypothesis is supported by a non-significant difference in the case of estimates of observed proportion vs. estimates of expected long-run proportion.

The main effect B, order of objective proportion estimates were based upon, yielded an F of 5.207 which for 1 and 420 degrees of freedom is significant beyond the .05 level. This significance shows that the mean error of 17.96 for estimates based upon the order square - triangle - triangle - square - triangle - square (C_2), is significantly larger than the mean error of 15.00 for estimates based on the order, triangle - square - square - triangle - square - triangle (C_1). The trial effect C failed to yield a significant F, F = 0.749 indicating that while there was a trend for the errors to decrease there was no significant difference between the trial mean errors (see Figure 2).

The interaction BXC, order x trials yielded an F of 11.668, which for 5 and 420 degrees of freedom is significant beyond the .01 level. Contained within this order x trials interaction is the effect due to the objective proportion that the estimates were based upon. This effect was pulled out and is shown in Table 3. The effect due to objective proportion yielded an F of 61.004 which for 1 and 420 degrees of freedom is significant beyond the .01 level. The interaction, objective proportion x trials, yielded an F of 2.233 which for 5 and 420 degrees of freedom is significant beyond the .05 level. The significance for objective proportion shows that the mean error of 20.48 for estimates based upon triangles (objective proportion of .67) is significantly larger than the mean error of 12.48 for estimates based upon squares (objective proportion of .22).

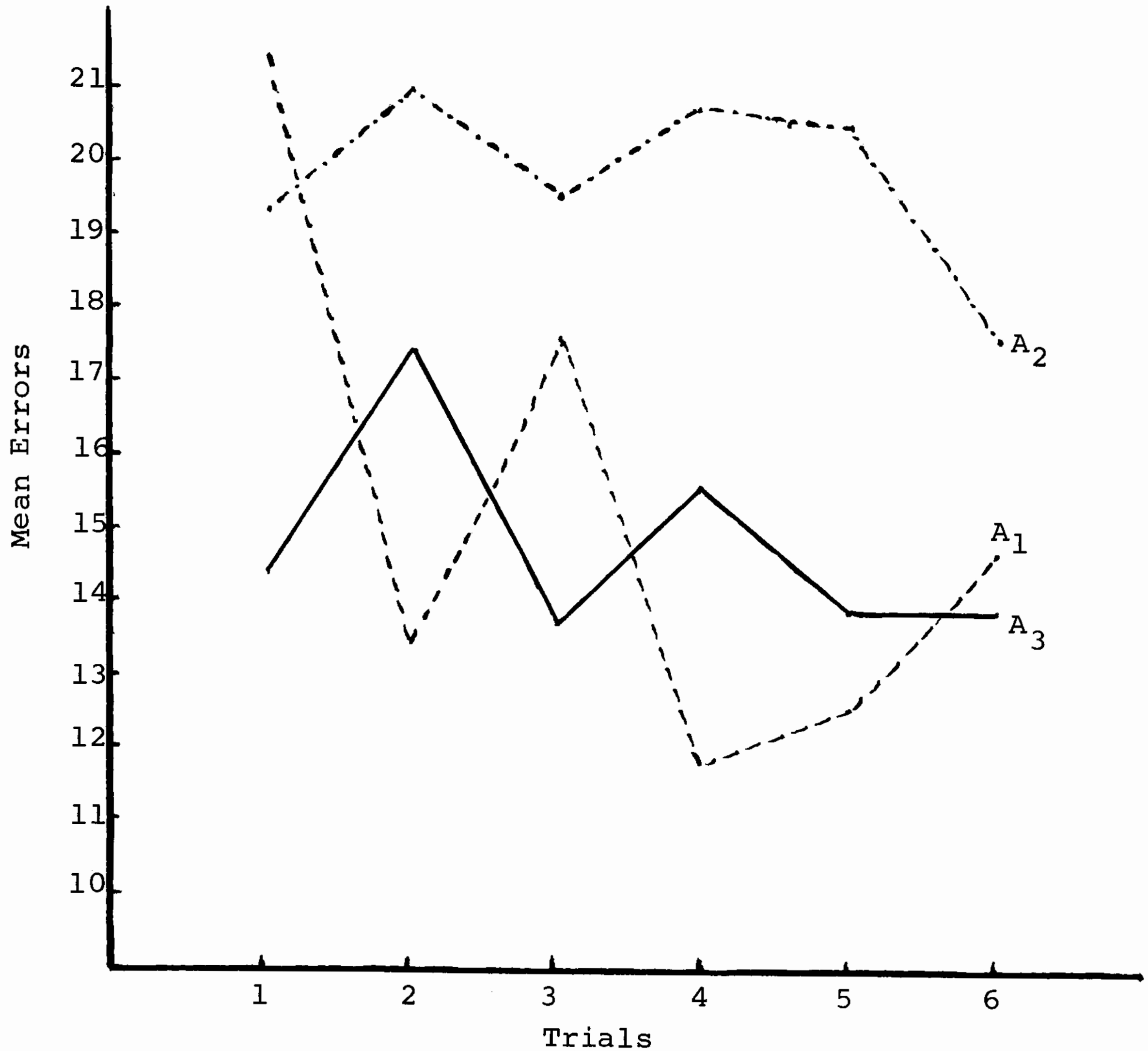


Figure 2. Means, averaged over levels of order and proportion for each of 6 trials.

DISCUSSION

The results of this study fail to reject the null hypothesis in one case and force rejection in two cases. A significant difference in mean error scores was found between estimates of observed proportion and estimates of subjective probability. This difference indicates that when using a direct psychophysical method, asking for an estimate of proportion is not the same as asking for an estimate of probability of occurrence. The significant difference between mean error scores of estimates of expected long-run proportion and mean error scores of estimates of probability of occurrence suggests that these are also different tasks. The non-significant difference between estimates of observed proportion and estimates of expected long-run proportion suggests that they are similar tasks. That is, asking a person to project an estimate into the future does not interfere with that estimate.

In several investigations of subjective probability, Edwards (1961) and Luce, et al. (1965), the investigators substituted estimates of observed proportion for estimates of subjective probability. This suggests that these experimenters believed these two tasks to be the same. It was felt by this investigator that the type of estimate asked for might produce different estimates. This study was designed to compare estimates of subjective probability with estimates of observed proportion when those estimates were based upon the same stimuli. It was hypothesized that asking for estimates of subjective probability would yield results similar to asking for estimates of observed proportion. The results of this investigation show that the methods used in it to obtain estimates do differ. When asked to estimate subjective probability Ss made significantly greater errors than when asked to make estimates of observed proportion or expected long-run proportion.

There are several possible explanations for these results. First, estimates of subjective probability may not be simple functions of estimates of observed proportion. It seems quite possible that while mathematical probability is a direct function of relative frequency or proportion, subjective probability may not be a simple function of <u>observed</u> proportion. Those investigations of the applicability of the additivity theorm to subjective probability (Holmberg, 1964) would seem to support this position. An additional analysis of the data from both raw scores and error scores points to further considerations. When comparing mean raw scores for each of the estimates it appears that all methods of asking for estimates produce similar results (see Table 4). For each type of estimate Ss underestimated the higher objective proportion (.67) and overestimated the lower objective proportion (.22). These results are in agreement with Preston and Baratta (1948) and Shuford (1959). It would appear that while Ss are not particularly accurate estimaters of these objective proportions they do give similar estimates when asked to estimate subjective probability based upon the same stimuli. That is, the results are similar regardless of type of estimate asked for.

However, when estimates are converted into absolute error scores the results differ. The results of this study show that when measured in error scores there is significantly greater amount of error for estimates of subjective probability than for estimates of observed proportion or estimates of expected long-run proportion. This suggests that comparing mean raw scores masks some real differences in estimates. This greater amount of error or variability in estimates of subjective probability could indicate that asking for estimate of subjective probability is a somewhat different task than asking for estimates of either observed proportion or expected long-run proportion.

Although the possibility remains that the differences between estimates is due to basic differences in the tasks involved, a second possibility exists. While an attempt was made to balance the instructions and the way estimates were asked for, it remains quite possible that certain key terminology was not equal for all treatments. In comparing the instructions for estimates of expected long-run proportion and for estimates of observed proportion the key terminology is "proportion." On the other hand, in reviewing the instructions for estimates of subjective probability, the key terminology is "probability." It may be that the tasks are the same, but that this difference in basic terminology caused the significant differences in estimates. It was assumed that following the elaborate instructions concerning probability, proportion, and expected long-run proportion that Ss could handle the concepts of proportion and probability. It seems reasonable to believe that probability is not a simple nor necessarily a familiar concept and may in fact seriously interfere with the measurement of subjective probability when asking for direct estimates.

There is the possibility that Ss are accurate in estimating subjective probability but that asking for a direct report of that estimate may compound the measure of subjective probability. This seems to be a worthy basis for suggesting that further studies be conducted investigating different methods of asking for estimates of subjective probability including indirect as well as direct psychophysical techniques.

In light of the significant differences between estimates found in this study, a number of considerations for future studies are suggested:

1. Estimates of relative frequency will not substitute for estimates of subjective probability where direct psychophysical judgment methods are used.

2. Future studies need to be centered around

investigating different methods of asking for direct estimates of subjective probability.

3. Consideration needs to be given to the idea that subjective probability and mathematical probability, while similar in appearance, may in fact be very different constructs.

While this study was designed to specifically investigate the relationship between types of estimates asked for, a number of other relationships were included both as controls and as a matter of interest. Ss were asked to give estimates based upon two different objective proportions. The results indicate that when Ss were asked to give estimates based upon an objective proportion of .67 they made significantly more errors than when estimates were based upon an objective proportion of .22. This objective proportion of .22 corresponded to a relative frequency of 16. Because each of the five (5) slides in each trial represented a different matrix, and the units of each matrix were randomized separately, it is assumed that Ss were unable to count and arrive at a relative frequency of 16. It is very possible that estimating a relative frequency of 16 is a much easier task than estimating a relative frequency of 48 when those two relative frequencies are not complimentary.

A possible explanation for this significant difference is found in a discussion of the results of another experiment (Shuford, 1961). The possibility was raised that perception of proportions has some of the characteristics of a sampling process. The results showed that the average variance of subjects increased as the proportion approached 50 per cent. A curve that represented the average variance vs. proportion was bimodal somewhere between 30 per cent and 40 per cent and between 60 per cent and 70 per cent, dropping somewhat at 50 per cent. Shuford (1961) suggests that this may indicate subjects base their estimates on a sample of elements in the matrix and further that subjects might tend to fixate on areas composed of clusters of elements. When objective proportions are at the ends of the scale fixating on clusters would lead subjects to closer estimates of proportions because at these proportions there is a greater chance of clustering and thus a cluster would more than likely be representative. When estimating proportions on either side of 50 per cent fixating on clusters would cause an increase in variance among estimates and would show up as a bimodal curve. Since an objective proportion of .22 is closer to the end of the scale than one of .67, this would seem to be a plausible explanation for the significant difference between these two objective proportions.

The significantly greater mean error for estimates based upon the sequence .67, .22, .22, .67, .22, 67 vs. the mean error for estimates based upon the sequence .22, .67, .67, .22, .67, .22 is perhaps the most surprising result. A possible explanation is based upon the previous discussion. If we assume that the task of making an estimate based upon a proportion close to the end of a scale is an easier task than that of making an estimate based upon a proportion near the center of that scale, then, the possibility exists that the initial trial (first estimate) may produce a set that lasts the duration of the sequence of estimates. Having an easier task to begin with may make subsequent tasks relatively easier, while having a hard task first may make subsequent tasks relatively harder. This type of relationship would result in greater errors for orders beginning with the target objective proportion of .67.

SUMMARY

The purpose of this study was to extend the investigation of subjective probability by exploring three methods of asking for estimates. The study included an examination of two methods of asking for estimates of subjective probability and one method of asking for estimates of relative frequency. The study utilized 90 college students viewing projected matrices, each matrix contained 72 stimulus elements made up of 48 triangles, 16 squares, and 8 circles. Six trials were given, on three Ss were asked to give an estimate based upon triangles and on three Ss were asked to give an estimate related to squares. An analysis of variance of the data in a randomized group design with repeat measures yielded two significant Fs for main effects and one significant F for an interaction.

REFERENCES

Becker, G. M. Decision making: objective measures of subjective probability and utility. Psych. Rev., 1962, 69, 136-148.

Coombs, C. H. and Beardslee, D. On decision-making under uncertainty, in Thrall, R. M., Coombs, C. H., and Davis, R. L. (Eds.) Decision Processes. New York: Wiley & Sons, 1954.

Crawford, J. J. The effect of certain stimulus characteristics upon quantitative inductive judgments. Unpublished dissertation, University of Washington, 1963.

Davidson, Suppes, Siegal

de Finetti, B. La prevision: ses lois logiques, seo sources subjectives. Ann. Inst. Poincare', 1937, 7, 1-68. English translation in Kyburg, H. E. Jr. & Smokler, H.E. (Eds.) Studies in subjective probability. New York: Wiley, 1964.

Edwards, A. L. Experimental design in psychological research. New York: Holt, Rinehart, & Winston, 1964.

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

Edwards, W. The theory of decision making. Psychol. Bull., 51, 380-411.

Edwards, W. Utility, subjective probability, their interactions, and variance preferences. J. conflict Resolut., 1962, 6, 42-51.

Edwards, W., Lindman, H., & Savage, L. J. Baysisian Statistical Inference for psychological research Psychological Review, 1963, 70, 193-242.

Edwards, W. Subjective probabilities inferred from decisions. Psychol. Rev., 1962, 69, 109-135.

Edwards, W. Reward probability, amount, and information as determiners of sequential two-alternative decisions. J. Exp. Psychol., 52, 177-188.

Edwards, W. Behavioral decision theory. Annu. Rev. Psychol., 1961, 12, 473-498.

Edwards, W. The prediction of decisions among bets. J. Exp. Psychol., 50, 201-214.

Edwards, W. Probability-preferences in gambling. Ann. J. Psychol., 1953, 66, 349-364.

Erlick, D. E. Judgment of the relative frequency of a sequential series of two events. J. Exp. Psychol., 1961, 62, 105-112

Feather, N. T. Subjective probability and decision under uncertainty. Psychol., Rev., 66, 150-164.

Holmberg, B. K. Application of the additivity theorem to subjective probability. Unpublished Master's Thesis. Central Washington State College, 1964.

Irwin, F. W. Stated expectations as functions at probability and desirability of outcome. J. Personality, 1953, 21, 329-35.

Luce, D. R., Bush, R. R., and Galanter, E. Handbook of Mathematical Psychology, Vol. III. New York: Wiley & Sons, 1965.

Mosteller, F., Rourke, R. E. K., and Tomas, G. B., Jr. Probability with statistical application. Reading: Addison-Wesley, 1964.

Neimark, E. D. & Shuford, E. H. Comparison of predictions and estimates in probability learning situation. J. Exp. Psychol., 1959, 57, 294-298.

Peterson, C. R., Schneider, R. J., & Miller, A. J. Sample size and the revision of subjective probabilities. J. Exp. Psychol., 1964, 69, 522-527.

Pruitt, D. G. & Hoge, R. D. Strength of the relationship between the value of an event and its subjective probability as a function of method of measurement. J. Exp. Psychol., 1965, 69, 483-489.

Savage, L. J. The foundations of statistics. New York: Wiley & Sons, 1954.

Shuford, E. H. A comparison of subjective probability for elementary and compound events. U. N.C. Psychom. Lab. Rep., 1959, No. 20.

Shuford, E. H. Percentage estimation of proportion as a function of element type, exposure time, and task. J. Exp. Psychol., 1961, 61, 430-436.

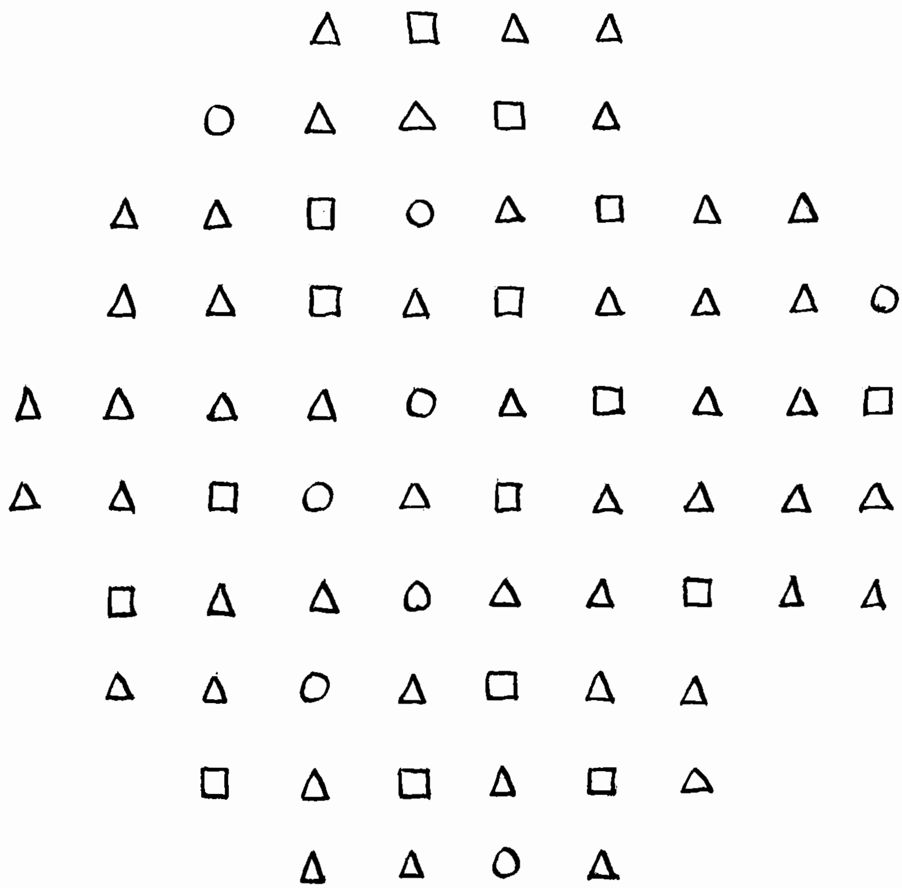
Stevens, S. S. & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. J. Exp. Psychol., 1957, 54, 377-409.

Thrall, R. M., Coombs, C. H., & Davis, R. L. Decision processes New York: Wiley & Sons, 1954.

A P P E N D I X

APPENDIX A

ONE OF FIVE DIFFERENCE MATRICES USED AS STIMULI



APPENDIX B

Instructions for each group based upon factor A, type of estimate asked for:

A_1--Observed proportion.

The task you are about to perform deals with the concept of proportion. When we speak of proportion we are referring to the mathematical relationship between events, objects, or things as determined by their relative number. In this way a proportion may be used to express the mathematical relationship between the members of a sub-class or sub-group and a larger class or group containing that sub-group. For example, if we had 100 things in a group and these were divided into four sub-groups containing 12, 35, 50, and 3 elements, then the relationship between each of these sub-groups and the total group could be expressed as a proportion. The most common way to express proportion is as a fraction. In the example given the proportions would be 12/100, 35/100, 50/100, and 3/100. When the proportion is small the numerator is much smaller than the denominator and when the proportion is large the numerator is almost as large as the denominator.

A_2--Probability of Occurrence of Next Event

The task that you are about to perform deals with the concept of probability. When we speak of probability we are referring to the mathematical relationship between events, objects, or things as determined by what we think is the relative likelihood of their occurrence. In this way a probability may be used to express the mathematical relationship between the likelihood of occurrence of members of one class and members of other classes. For example, if we expected a group to have 100 things in it and expected that group to be divided into four different sub-groups containing

12, 35, 50, and 3 elements, then the relationships between the likelihood of occurrence of a member of one sub-group as compared to the likelihood of occurrence of a member of members of other groups could be expressed as a probability. The most common way to express probability is as a fraction. In the example given the probabilities would be 12/100, 35/100, 50/100, and 3/100. When the probability is small, the numerator is much smaller than the denominator and when the probability is large the numerator is almost as large as the denominator.

A_3--Expected Long Run Proportion

The task that you are about to perform deals with the concept of expected proportion. When we speak of expected proportion we are referring to the mathematical relationship between events, objects, or things as determined by what we think the relative number of each will be. In this way, an expected proportion may be used to express the mathematical relationship between the members of a sub-class or sub-group and a larger group or class containing that sub-group. For example, if we expected a group to have 100 elements in it, and we expected that group to be divided into four sub-groups, containing 12, 35, 50, and 3 elements, then the relationship between each of these sub-groups and the total group could be expressed as an expected proportion. The most common way to express expected proportion is as a fraction. In the example given the expected proportions would be 12/100, 35/100, 50/100, and 3/100. When the expected proportion is small, the numerator is much smaller than the denominator and when the expected proportion is large the numerator is almost as large as the denominator.

APPENDIX C

Method of asking for estimates:

A_1 (B_1 and B_2) Observed proportion for objective proportions of .22 and .67.

Considering the sets of geometric figures you have just seen, what proportion of them were squares (.22)? ..... were triangles (.67)?

A_2 (B_1 and B_2) Probability of Occurrence of Next Event Based upon objective proportions of .22 and .67.

If I were to show a single geometric figure to you, based on the slides you have just seen, what would you expect the probability to be that the next slide shown would be a square (.22)? ..... a triangle (.67)?

A_3 (B_1 and B_2) Expected Long Run Proportion for objective proportions of .22 and .67.

If I were to continue to show this set of slides to you, what in the long run would you expect the proportion of squares (.22) to be? ..... of triangles (.67) to be?