DEVELOPMENT OF THE

SPATIAL JUDGMENT EXPERIMENTAL

TASK

by

James Curtis Moore, Jr.

Technical Report No. 15

July, 1965

DEVELOPMENT OF THE SPATIAL JUDGMENT EXPERIMENTAL TASK*

The experimental task to be discussed in this report was developed for use in the author's dissertation research. Its particular characteristics are, however, general enough to make the task applicable to other research settings.

As Thomas Conner (1964) has noted, the design of laboratory experiments to test formal models of social processes confronts the investigator with the responsibility of analyzing in detail the nature of the activities experimental subjects are asked to engage in. This analysis must include pre-experimental tests to determine if the behavior of subjects engaged in these activities in the absence of experimental manipulations can be characterized in a theoretically specified manner. In brief, the objective is to construct an experimental task with characteristics such that the task itself does not interfere with the process of experimental interest. In the absence of such ideal attainment, one must at least be able to specify in some detail the impact that the task itself has upon this process of interest.

At the time the author's dissertation research was being planned, the Laboratory for Social Research had at its command three experimental tasks which had been extensively pretested. In one or more respects, each of these tasks had failed to meet requirements which were deemed desirable for

---

*I must express my gratitude to Robert Muzzy and Gordon Lewis for their many helpful suggestions relating to the construction of this task; to Thomas Conner for his help in collecting the data and to both Conner and Hans Lee for their suggestions regarding analysis of the data; to Gerald King for his time in developing computer programs; and to Professors Bernard Cohen and Morris Zelditch, Jr., for their encouragement and willingness to provide financial assistance.

experimental work. But just as importantly, these ambiguous tasks were not capable of being later expanded into veridical or "near veridical" tasks. Extensions of the author's dissertation research call for such a task, and this property seems to be, in general, desirable. For these reasons, the development of an entirely new task was undertaken. The "perceptual" character of the Spatial Judgment Task is the result of an attempt to avoid certain problems which became apparent during the testing of the three earlier tasks.(See Conner, 1964). It was felt that these problems were most likely derivative of verbal associations which subjects made with the stimuli components.

The basic requirement of the task was that it be composed of a series of discrete decision stages. At each decision stage subjects are required to evaluate and choose between two mutually exclusive alternatives. The research interests demanded that each discrete choice be ambiguous to the subject in the sense that standards for making a choice be ill-defined or unaccessible. This "ambiguity" property implies that, consistent over individuals and decision stages, there should be no bias towar one choice alternative over the other. It was further required that the investigator be able to convince subjects that there was a specific ability associated with the making of these choices. Finally, the choices should be independent of one another. That is, the choice at each stage should be uninfluenced by the choice at any past stage.

After a good deal of preliminary exploration, a "checkerboard" pattern of black and white rectangles was deemed the most appropriate stimulus. Five original stimuli plates were constructed, using black poster color on heavy white posterboard. Each plate measured 7 and 1/2 inches in height and 15 inches in length, with the area divided into one hundred small rectangles each measuring three-quarters of an inch by one and one-half inches. By random selection fifty of these small rectangles on each plate were colored black, with the remaining fifty rectangles left white. The random selection of rectangles to be colored was tempered by the necessity to avoid striking patterns within the grid. Each of the five finished plates was then mounted on grey background paper and photographed with panchromatic plus-x 35mm. film. Lighting and development were carefully controlled to insure constancy of stimuli.

Each stimulus plate yielded a set of eight variant stimuli: the plate photographed in an "upright" or "first" position and in a 180-degree rotated position; a color reversal of the plate photographed in these two positions; an "A" plate which was prepared by rotating the original 90 degrees to the right and then reconstructing the rectangle, again photographing in the two mentioned positions; and a color reversal of the "A" plate, again photographed in these two positions. Despite the intimate construction relationships between the stimuli in any set of eight, inspection indicates that each stimulus has a substantially different character from any other stimulus in the set. In addition, since the stimuli were prepared as 35mm. transparencies for projection, each can be placed into a projector in two positions, arbitrarily called "front mounted" and "back mounted". This therefore yields, from one original construction, sixteen essentially different stimuli. A photographic print of a single stimulus is presented below.

The pretesting of the stimuli was carried out in seven phases. The first two phases were "field" operations with the stimuli being administered to large groups of individuals at a single time. One hundred and forty-eight individuals were employed in this part of the pretesting. The final five phases of the pretesting were carried out under standard laboratory conditions at the Laboratory for Social Research. In the laboratory individuals responded to the stimuli by pressing choice buttons which lit bulbs on the experimenter's control panel, and the choices were subsequently recorded by the experimenter. In the field pretests, subjects responded to the stimuli by circling responses on an answer form. One hundred and twenty individuals were employed for the laboratory phases.

Phase 1: Subjects for this phase were recruited at Stanford by placing an advertisement in the daily newspaper. The pretest was advertised as a simple perception study, and each participant received one dollar. Thirty-five subjects were recruited in this manner. After the "test", subjects were informed that we were interested in locating the sources of perceptual biases in these stimuli, and that no scores could be made available to them. Subjects were satisfied with this situation after we had explained that such scores would be meaningless due to the fact that relationships between this "spatial judgement ability" and other abilities had not been specified. In the field phases of the pretesting, subjects were not informed that each slide was 50% black and 50% white, although they were so informed in the laboratory phases of the study, inasmuch as we could talk to each individual in some detail regarding our research objectives. A copy of the instructions used during this first phase is included in the appendix to this report. The same instructions were used in phase two.

This first pretest consisted of two demonstration stimuli (the first being 54% white and the second being 54% black in construction), followed by forty-six stimuli 50% white and 50% black in construction. In this set of forty-six, two different stimuli were presented four times each, although not in succession. The rationale for this will become apparent. The demonstration stimuli were constructed as stated in order that we could leave a stimulus on the screen for a long period of time while explaining the task to subjects, and in order to lend validity to the requirements which the subjects were told to meet as they worked on this task. Responses to the two demonstration stimuli are of course not included in the analysis.

The desired ambiguity property implies that the probability of a white response (or a black response) has the <u>a priori</u> probability value of .50. The best estimate of p is the mean proportion of white (or black) choices over all subjects and all trials. The assumption of independent trials means that the proportion of white or black choices, controlling for the immediately previous choice, is also p. This assumption can be evaluated in terms of an "aggregate transition matrix" in which the columns of the matrix give the proportion of each choice on trial n + 1 given the particular choice on trial n. Independence means that the entries in a given column are identical. Thus, an aggregate matrix indicating independence of trials will appear as follows.

		n + 1		
		Black	White	
n	Black	α	$1 - \alpha$	α
	White	α	$1 - \alpha$	$1 - \alpha$
		α	$1 - \alpha$	N

The X^2 test for independence can be utilized as a rough indication of whether column entries are in fact identical. However, with large N, a X^2 with a probability value of .001 may indicate only the slightest departure from independence. The X^2 values themselves should therefore not be taken as the absolute guide. Tempered subjective judgments are not inappropriate at this point.

The "aggregate transition matrix" is of course an end product of the combination of all the one-step transition matrices. The examination of the aggregate matrix to test independence assumes that all of the one-step matrices are statistically equivalent or homogeneous. The necessity for this can be seen clearly by considering a one-step transition matrix which demonstrates perfect major-diagonal dependence and a one-step transition matrix which demonstrates perfect minor-diagonal dependence. The conjunction of these two matrices produces an aggregate matrix demonstrating independence. If there is reason to doubt this homogeneity assumption, X^2 can be computed for each of the one-step transition matrices. These values are then summed, and this sum is distributed as X^2 with degrees of freedom equal to the sum of degrees of freedom of each of the separate X^2 values. (Goodman, 1962) This computation is quite tedious, and by that fact alone the examination of the aggregate matrix is the preferable procedure if the assumption of homogeneity is reasonable. There are no theoretical grounds leading one to question this assumption, but in the analysis of the data from phase two, independence will be examined by this latter procedure.

Returning now to the data from phase one, and letting p equal the probability of a white response, the following aggregate matrix was obtained.

		n + 1 Black	n + 1 White	
n	Black	.42	.58	.39
n	White	.38	.62	.61
		.40	.60	1575

$X^2 = 3.211$ $\quad .10 > p > .05$

Independence appears clearly indicated in this matrix, but it is also clear that there is a substantial inflation of the probability of a white response. The probability of securing a sample of this size with $P_s = .60$ when in fact $P_u = .50$ is less than .001. The first suspicion was that this inflation was an artifact due to particular conditions existing in this "field" test; viz., the use of a glass bead screen, the relatively great distance of the subject from the screen, and closely related to this, the possibility of subject eyesight problems. As will become apparent upon examination of data from other phases of the pretest, this speculation is not correct. There is indeed a substantial inflation of the probability of a white response which is a function of the stimuli themselves. We will return to this point later.

Further examination of the task in terms of the criteria that have been set forth requires an examination of the distribution of individual subjects in terms of the total number of white responses given and an examination of the distribution of individual stimuli in terms of the total number of white responses which they elicit. Theoretically, these should be binomial distributions.

The distribution of subjects for this phase of the pretest was quite encouraging. Given p = .60, there were no areas of substantial deviation from the theoretical binomial. X^2 was found to be 7.064 with 4 df, yielding a probability level between .20 and .10. The distribution of stimuli, however, indicated substantial deviations from the theoretical distribution. It is not necessary to present these graphs at this time.

In an attempt to locate the source of the latter deviation, the entire stimuli set was carefully examined. One striking feature was noted. In some of the stimuli a vertical line appeared as a result of the non-overlapping conjuction of small black and white rectangles. In some cases, the vertical line divided the total stimulus into a large subrectangle and a small subrectangle. In other cases the vertical line divided the total stimulus into two equal squares. It was hypothesized that in those cases where the vertical line resulted in a division of the total stimulus into two subrectangles and where the large subrectangle was not 50% white and 50% black, there would be a response bias in the direction of the predominant color in the large subrectangle. It was suspected that when faced with this "ambiguous" task, the subjects' attention would be drawn to the large subrectangle and they would employ their judgment of this subrectangle as the "best estimate" of the total rectangle. In those cases where the vertical line divided the total stimulus into two equal squares, it was hypothesized that a "reading order" effect would be observed. The expectation was that the left-hand side or top would bias the response.[1]

Those stimuli characterized by the off-center vertical line were variations of a single original plate. The plate was #2, and the large subrectangle was 54% white in construction. The following results were obtained:

	Probability of White Response
Plate 2 in first position	.942
Plate 2 rotated 180 degrees	.857

Now notice what happens when the large subrectangle becomes 54% black:

Plate 2 color reversed	.285
Plate 2 color reversed and rotated 180 degrees	.257

The effect appears unequivocal. It will be recalled that the "A" plate was simply a 90 degree rotation and reconstruction of the original. The vertical line is thus converted into a horizontal line, still dividing the large rectangle into two unequal subrectangles. Even here, the effect is striking.

Plate 2A in first position	.771
Plate 2A rotated 180 degrees	.714
Plate 2A color reversed	.314
Plate 2A color reversed and rotated 180 degrees	.342

[1] Mr. Gordon Lewis suggested this possibility.

Examination of the data for those stimuli which possess a vertical line dividing the total stimulus into two squares gives some support to the reading order hypothesis as well, though not nearly so striking as the above. Again, eight such stimuli were detected, and they were all variants of a single original construction. Beyond any doubt, a cardinal principle for the construction of these stimuli has been found. In the phase two pretest, these biased sixteen stimuli were removed from the set.

Before presenting the data from phase two, however, consideration must be given to the earlier subjective judgment that the stimuli constructed as variations of a single original plate are perceptually different stimuli. In this first phase, all stimuli were "front mounted" into the projector. In phase two, both "front" and "back" mountings were employed. For the present, we will evaluate only the hypothesis of independence between a given stimulus presented in the "first" or "upright" position and that same stimulus presented in the 180-degree rotated position. X^2 calculations were performed upon four-fold tables constructed as follows.

		Stimulus "x" in 180 degree rotated position: Black	Stimulus "x" in 180 degree rotated position: White
Stimulus "x" in first position	Black	Number of Subjects	Number of Subjects
	White	Number of Subjects	Number of Subjects

35 (in all cases)

Utilizing only those stimuli free of the vertical line bias, 11 such tables were constructed. Of the eleven X^2 values, nine were associated with a p value of .30 or greater, one was associated with a p value between .20 and .10, and one was associated with a p value between .05 and .02. There is clearly insufficient evidence for rejecting the hypothesis that these stimuli are perceptually different, despite their intimate construction relations. Even more importantly, the same stimulus presented in precisely the same position on each of four non-successive trials elicits independent responses. It will be recalled that two of the stimuli were presented four times each. One of these stimuli demonstrated the vertical line bias, and will not be considered here. The other stimulus, however, gave the following results.

Stimulus 1 in First Position

(assume $p_w = .60$)

Response Pattern	p	Expected No. S's for response pattern	Observed
4 Black	.0256	0.896	2.000
3 Black 1 White	.1536	5.376	8.000
2 Black 2 White	.3456	12.096	10.000
1 Black 3 White	.3456	12.096	9.000
4 White	.1296	4.536	6.000
		35.000	35.000

$X^2 = 2.217$ $\quad .70 > p > .50$

It appears that unless there is a particularly striking feature within these stimuli(such as the vertical line), even the <u>same</u> stimulus fails to elicit recognition. This fact of course implies great flexibility in the possible ways to combine these stimuli into sets for experimental use.[2]

<u>Phase 2</u>: Subjects for this phase were students at San Francisco City College and the University of San Francisco whose instructors had volunteered a class period to facilitate the pretesting of this task. Subjects were not paid in this phase of the work. Instructions and the explanation given were the same as in phase one. One hundred and thirteen subjects were employed. The stimuli presented were those from phase one which demonstrated no vertical line features. Each was mounted as in phase one, with the addition of a "front" and "back" mounting into the projector. As before, the two described demonstration stimuli were presented, followed by forty-six "ambiguous" stimuli.

The aggregate transition matrix derived from this phase is given below.

$$
\begin{array}{ccccc}
 & & \multicolumn{2}{c}{n+1} & \\
 & & \text{Black} & \text{White} & \\
n & \text{Black} & .45 & .55 & .44 \\
 & \text{White} & .43 & .57 & .56 \\
 & & .44 & .56 & 5085
\end{array}
$$

$$\chi^2 = 1.486 \qquad .30 > p > .20$$

---

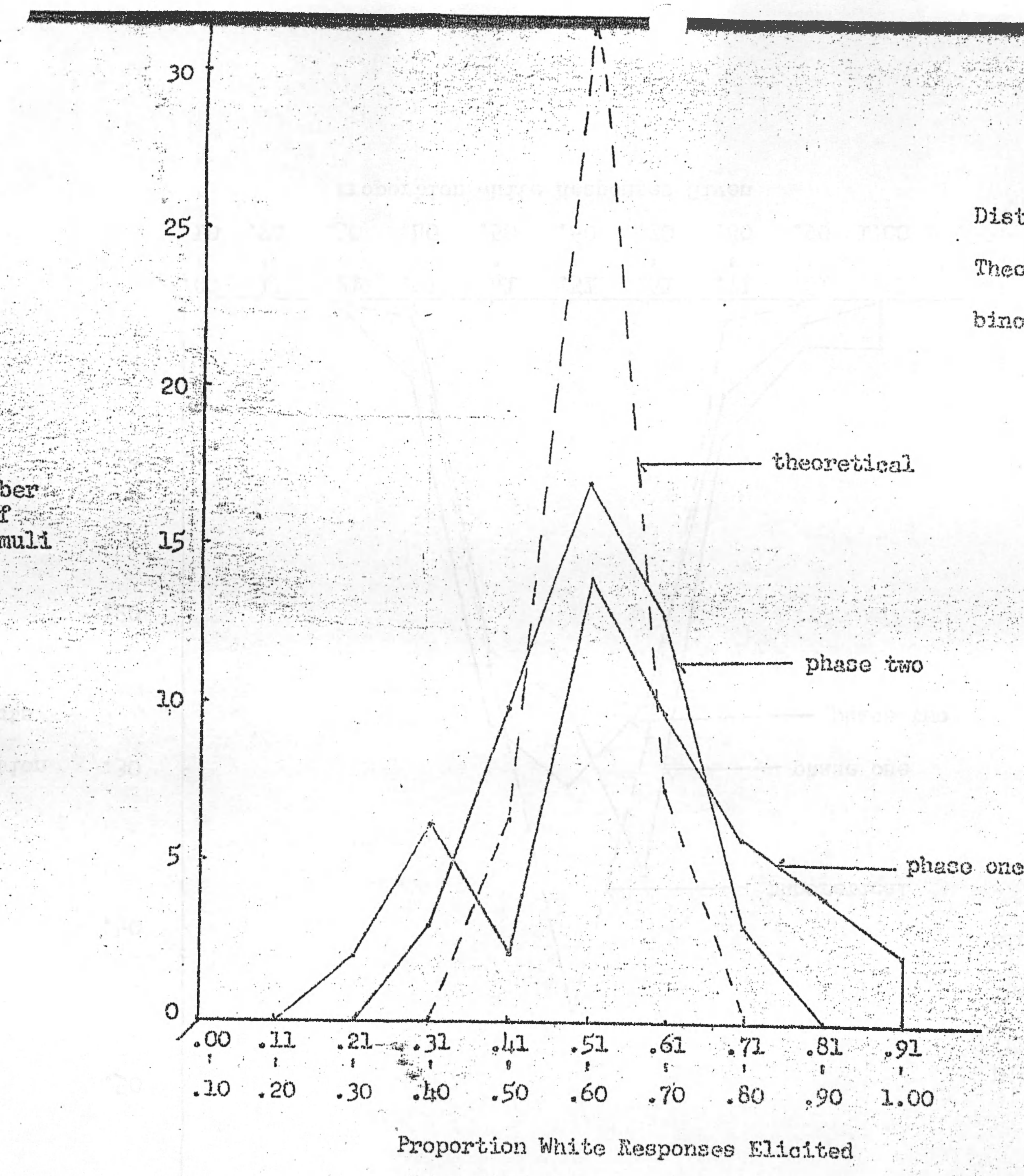
[2]Before proceeding, it should be reported that in this phase one data and throughout all of the pretesting, there was no indication that any features of the data could be considered to be the result of either age or sex differences among the subjects.

Once again, we find that the probability of a white response is somewhat inflated. The probability of securing a sample of this size with $P_s = .56$ when in fact $P_u = .50$ is 0.10. At this point, the speculation was still that this inflation was simply an artifact of the conditions under which the test was administered. Independence of trials appears clearly indicated in the above matrix.

Mention has already been made of the necessity for assuming homogeneity of the one-step transition matrices which make up this aggregate matrix. For this phase, a separate X^2 was computed for each of the one-step matrices for both stimuli and subjects. For stimuli, the summed X^2 value was 44.29 with 44 df, yielding a probability value of .47 (one-tail). Comparing the stimuli composing the first half of the series with stimuli composing the second half of the series indicated no time trend. For subjects, the evaluation of the homogeneity assumption entails the calculation of one hundred and thirteen X^2 values. Due to the relatively small N for each of these calculations (45), the standard correction for continuity should be applied. This further complicates the calculations. Consequently, each X^2 was initially computed without the correction, and only those values greater than 3.84 ($p = .05$) were reexamined using the correction for continuity. Looking then at the distribution of X^2 values, we would expect 5% of these to exceed 3.84. In fact, 3.6% of the X^2 values exceeded this value. There thus appears to be no necessity to question the homogeneity assumption, and all further analysis is based solely upon the aggregate transition matrices.

Turning now to a consideration of the distribution of stimuli and subjects, the following graphs are presented. In addition to the theoretical and observed distributions for this phase two data, each graph also presents the distributions attained in phase one. It is quite apparent that substantial improvements have been effected simply by removing those stimuli with the vertical line feature. However, despite the general improvements, the fits are not particularly good. For the data from phase two, $X^2 = 24.32$ with 2 df for stimuli and 8.93 with 2 df for subjects. Probability levels are less than .001 and approximately .02, respectively. Again, X^2 values should not be taken as the absolute guide, and the improvement of fit and general shape of the distributions is subjectively quite encouraging.

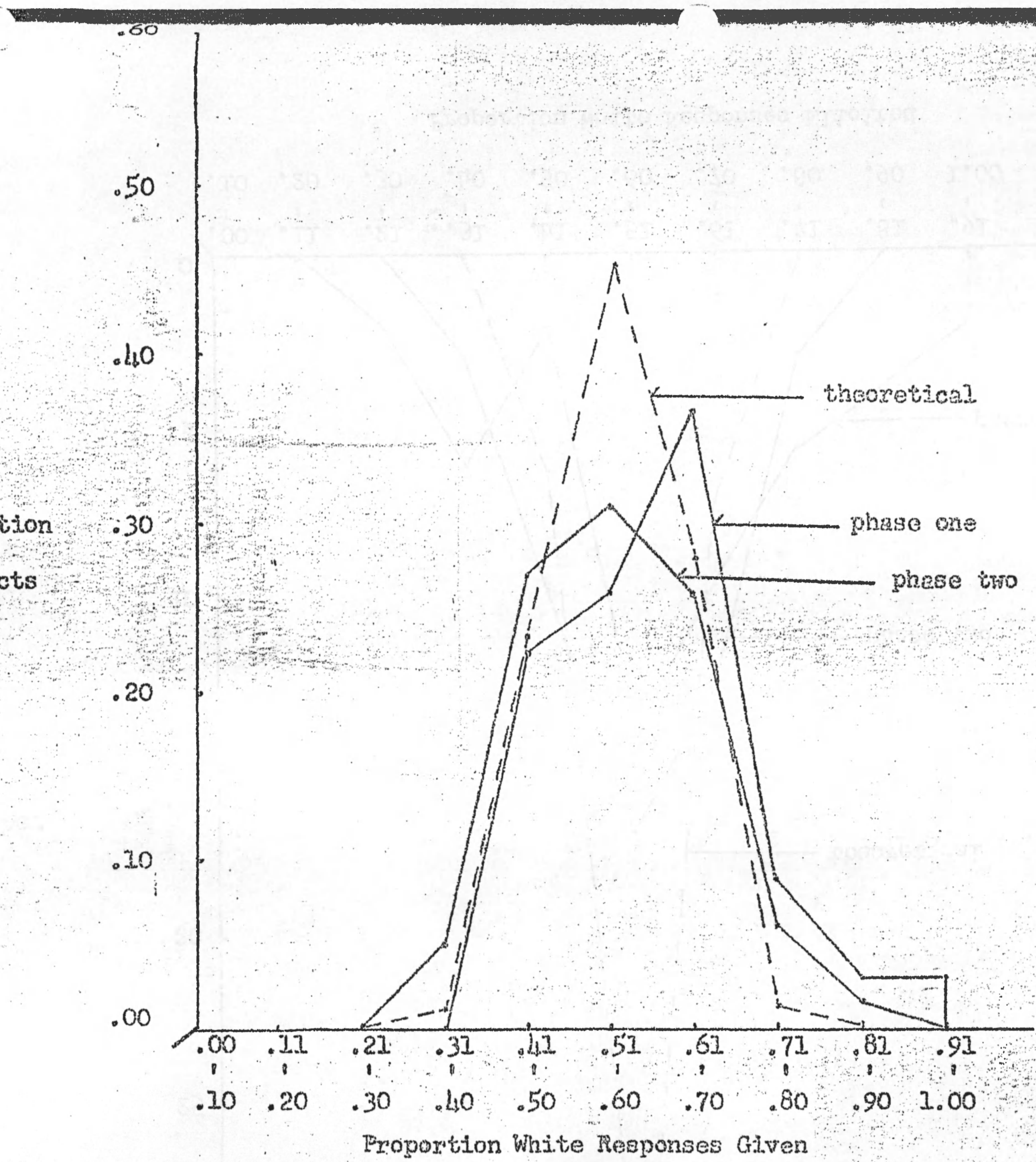
Close inspection of the two empirical curves suggests a striking platykurtosis, and this in turn suggests that both the stimuli and subject populations are composed of two rather distinct subpopulations. For stimuli, these two subpopulations would of course be "black-biased stimuli" and "white-biased stimuli". In the case of subjects, there could exist "black responders" and "white responders". The stimuli biases could be controlled in experimental work simply by presenting a randomized order to each experimental group. This solution, however, increases within-trial variability and must be compensated for by increasing the size of the subject samples utilized in experimentation. This is not always a feasible alternative. It is obviously desirable to find stimuli which will provide us with an acceptable distribution, and this problem has been pursued in a later phase of the pretest. For the moment, no



Distributions of Stimuli (N = 4

Theoretical distribution is the

binomial given p = .56.



Distribution of Subjects.
(Phase one N = 35; Phase two

Theoretical distribution is t

binomial given p = .56.

further discussion of this important problem is warranted.

If in fact the subjects could be classified into "white responders" and "black responders", this would pose extremely serious problems. Under such conditions, no pretense could be made that the stimuli set is composed of "ambiguous" stimuli. Over all subjects, a given stimulus might appear to be ambiguous, but each subject could have a good deal of subjective certainty with respect to his choices. That is, for a given individual, the choices might not appear ambiguous at all.

A careful examination of the subject distribution failed to substantiate such a division of the population. It appeared rather that there was an "ideal" distribution of subjects around .50 and another distribution around a probability of a white response of approximately .60. Thus, it appeared that there was a subset of the population which demonstrated the white bias with the remaining subjects responding as expected. These subsets appeared to be approximately equal in size and two separate theoretical distributions were thus computed; one with a p value of .50 and the other with a p value of .60. These two distributions were then added together, and the resulting distribution compared to the observed distribution. The fit was extremely good, yielding a X^2 value of 15.331 with 15 df and probability level between .50 and .30. This result fit quite nicely with the earlier speculation that the inflated probability of a white response was an artifact of the field conditions, and the conclusion was that those subjects at a particularly great distance from the screen were showing this white bias. Again the speculation proved to be untenable with later work. This point will be further discussed in the presentation of results from phase three.

Examination of independence between "front mounted" and "back mounted" stimuli again demonstrated the flexibility with which stimuli could be combined in a set. Of the twenty-two X^2 values computed, nineteen were less than 3.84 ($p = .05$). The remaining three displayed a p level of approximately .02. This is not substantial evidence for doubting the assumption of independence between these two mounting positions. The front and back mounting of these stimuli, conjoined with the previously discussed construction relationships, led several subjects to question whether certain of the slides had been repeated in the test. As a result of this fact, future instructions will incorporate a comment to the effect that "Certain people have felt that some of these slides have been repeated throughout the test. This is not the case. There are a very few basic patterns employed, with many slight variations on these. Spatial judgment ability is, in part, the ability to detect these slight differences." This statement was used with good results during the brief explanation given to the subjects participating in this phase of the work.

<u>Phase 3:</u>[3] Subjects for this phase of the pretest were Stanford males recruited by advertisement from the freshman and sophomore dormitories. Thirty-two subjects took part in this phase. Each was paid $1.50, and each received a rather detailed explanation of our research interests. Subjects arrived at the Laboratory for Social Research at a designated hour either singly or in pairs. The stimuli set employed was unchanged from phase two. Instructions utilized in

---

[3] I am grateful to Miss Lynn Roberts, Mr. Murray Webster and Mr. M. Hamit Fisek for the assistance they provided in this and subsequent phases of the pretesting.

this phase are included in the appendix to this report. Having made no changes in the stimuli set, the primary interest in this phase was whether the the inflated probability of a white response would be maintained under laboratory conditions.

The aggregate transition matrix obtained under these standard laboratory conditions is presented below.

		n + 1		
		Black	White	
n	Black	.38	.62	.38
	White	.38	.62	.62
		.38	.62	1440

Independence of trials is beyond question here. But it is just as clear that the inflated probability of a white response has been maintained, if not further inflated. This finding thus cannot be considered to be an artifact of the field conditions as was earlier suspected. Further, the fact that the subject population could be divided into an "ideal" population and a "white-biased" population in phase two may in itself have been an artifact resulting from the field conditions. It may well have been the case that in utilizing an answer form subjects could note the fact that they had a tendency to make long white response runs and in fact prevent this from occuring to any substantial degree. On a binary choice task, the possibility of all or most stimuli having the same "answer" is very likely not a tenable hypothesis for many subjects. In the laboratory, the subject presses a button to make his choice, and thus has no accurate record of his response pattern. Whatever the case, the inflation persists. The stimuli distribution for this phase shows no notable change from the distribution obtained in phase two. The subject distribution is a reasonably good fit to the theoretical distribution (given p = .62), with the exception that three of the thirty-two subjects gave such a large number of white responses (40 or more out of 46 possible) that they fall completely outside the upper tail of the theoretical distribution. No subjects make "too few" white responses. In the discussion with the subjects after the test, almost all mentioned that "white seemed to jump out". Many stated that they thought the first several slides were "white", but then concluded that "not all could be white". Examination of the data confirms the effect witnessed by these subjects. The first stimulus elicits approximately 85% white responses. This is true of the first stimulus in the earlier two phases as well, even though the first stimulus has been <u>different</u> for <u>each</u> of the three phases. The second and third stimuli also show a substantial white bias, and then the effect seems to dissipate. By plotting the probability of an alternation in the response from trial n to trial n + 1, the effect can be seen again. This alternations curve begins at a point much lower than is expected on the basis of the overall estimate of the probability of a black or white response, and does not stabilize at the expected value until after the first nine or ten trials.

The failure to account for the white bias in terms of conditions peculiar to the field administrations of the task of course led to a search for other reasonable hypotheses. The only reasonable hypothesis remaining was that the areas of the stimuli which projected "white" were far too intense. Since these stimuli are photographic transparencies, the white areas are simply areas of the film which allow the projected light to pass through. The black areas, on the other hand, are areas on the film which prohibit the light. It is therefore possible to decrease the intensity of the light striking the transparency and thereby decrease the intensity of the white areas without much affecting the black areas.

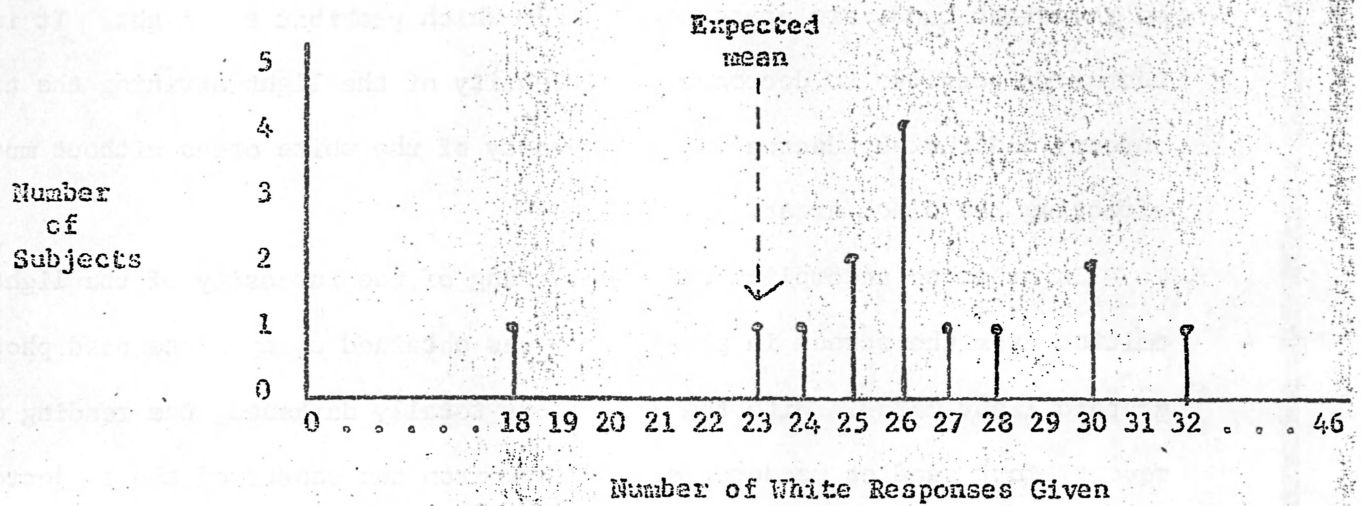
In order to accomplish this, a reading of the intensity of the light emitted from the screen in phase three was obtained using a standard photographic light meter. With the laboratory totally darkened, the reading was seventy foot candles measured at one foot from the center of the projector screen.[4] A Variac unit (a type of rheostat) was then wired into the projector power supply. This unit allows the operator to regulate the voltage which reaches the projector, and decreased voltage simply means decreased intensity.

In order to evaluate the efficacy of this solution, twenty-one male students from Menlo College were recruited to participate in the study. Again, each was paid $1.50 and each received a detailed explanation of our objectives. Instructions remained unchanged from phase three. Fourteen of these subjects constitute phase four, with the remaining seven constituting phase five.

---

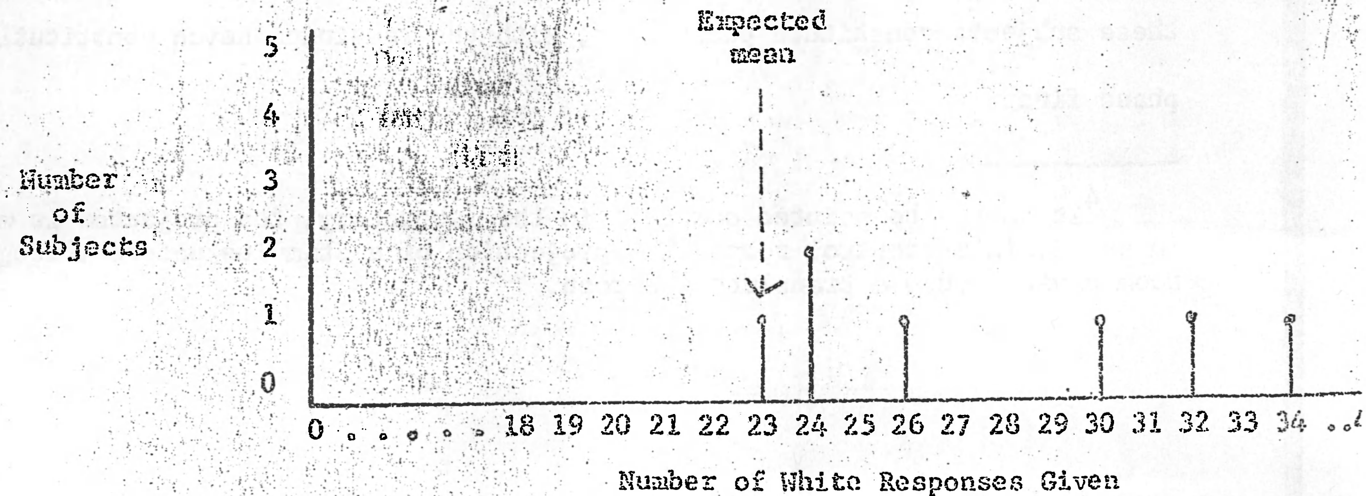
[4]It should be pointed out that in the laboratory, the projector is mounted in an adjoining control room. The projection of a stimulus passes through the room divider onto a translucent screen.

Phase 4: In this phase, the power reaching the projector was reduced by approximately one-half (to 65 volts). A light meter reading indicated eight foot candles, again measured at one foot from the center of the screen. This constitutes a striking reduction in intensity. The results can be summarized succinctly.



For the subjects in this phase, the estimated probability of a white response was .57.

Phase 5: In this phase, the power was further reduced to 55 volts. The light meter reading indicated six foot candles, measured as before. At this level, it appears that a limit has been reached. No further reduction can be tolerated. Even at this level, no change has been effected.



For these seven subjects, the estimated probability of a white response was .60.

Phase 6: Following the unsuccessful attempt to correct the white bias by controlling the intensity of projected light, it was reasoned that since the set of slides which were constructed 50% black and 50% white posed this problem, a set of the same basic stimuli constructed 52% black and 48% white might well solve the problem. Since later experimentation utilizing "near veridical" stimuli had already been planned, these latter stimuli were readily available.[5] This new set of stimuli was presented to thirty-five male subjects again recruited from Menlo College. All other conditions and procedures remained unchanged.

The aggregate transition matrix appears as follows.

		n + 1 Black	n + 1 White	
n	Black	.47	.53	.49
n	White	.53	.47	.51
		.50	.50	1575

$x^2 = 5.520$ $\qquad$ $.02 > p > .01$

---

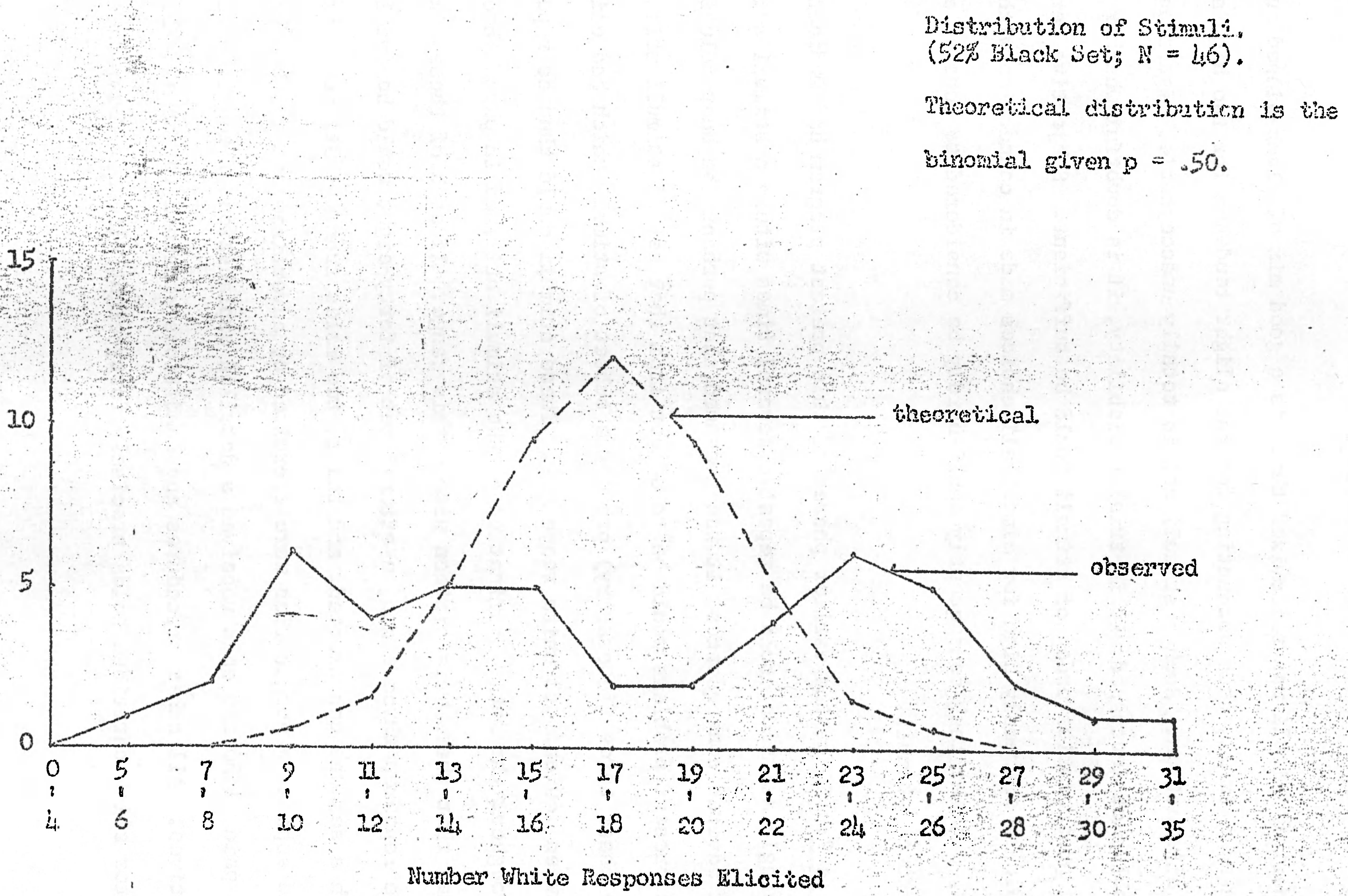
[5] In one photographing session, complete sets of these stimuli constructed 42% white, 44% white, 46% white, 48% white, 50% white, 52% white, 54% white, 56% white, and 58% white were filmed. Taking into account duplications of each stimulus, better than 1000 stimuli were filmed. Thanks are due to Mr. Marvin White for his conscientious work in processing these films.

On the basis of this matrix alone, one is tempted to conclude that the problem of the white bias has indeed been solved. The very slight departure from independence which is indicated would not be problematic. Of particular interest, however, is the fact that the slight dependence indicated is a minor-diagonal dependence. That is, there is some indication of too much alternation in the responses. This matrix remains consistent for the first half and second half of the data. Contrary to the phase three data, an alternations curve in the present case does not indicate a low rate of alternation initially. Rather, the curve begins slightly above .50 and remains constant, although slightly higher than the value expected on the basis of the estimated probabilities of a black and white response obtained from the above matrix.

The subject distribution obtained in the present phase was exceptionally gratifying. There were no substantial areas of deviation from the theoretical bionomial (given p = .50), and X^2 was found to be 4.605 with 6 df yielding a probability value between .70 and .50.

To this point, the results justified a cautious optimism, but examination of the stimuli distribution was something of a shock. Without much doubt, the stimuli set displays a sharp bifurcation into black-biased stimuli and white-biased stimuli. Approximately half of the stimuli would appear to fall into each category. This distribution is presented on the following page.

The nature of the stimuli set accounts for the earlier falsely encouraging results. Since black-biased stimuli are roughly equaled by white-biased stimuli, the subject distribution has been artifically improved and since the initial


Distribution of Stimuli.
(52% Black Set; N = 46).
Theoretical distribution is the
binomial given p = .50.
15
10
5
0
theoretical
observed
0 - 4
5 - 6
7 - 8
9 - 10
11 - 12
13 - 14
15 - 16
17 - 18
19 - 20
21 - 22
23 - 24
25 - 26
27 - 28
29 - 30
31 - 35
Number White Responses Elicited

order of stimuli was randomized there is a good mix of black-biased and white-biased stimuli, thus accounting for the slight tendency for too high a rate of alternation. Such a stimuli set is totally unacceptable. Ignoring the problem of increased within-trial variability, it is doubtful whether randomizing the presentation order of stimuli would be sufficient. Subjective certainty could be extremely high for almost all choices made in connection with this set, and the task consequently could hardly be considered to consist of "ambiguous" stimuli.

Phase 7: Since the earlier phases of this pretest program had indicated that a single stimulus could be repeated several times within a stimuli set without adverse consequences, it was reasoned that perhaps the best solution to the present problems would be to select from that set of stimuli utilized in the phase three (laboratory) pretest a subset of stimuli which had elicited between .40 and .60 white responses. These stimuli could then be repeated throughout the set. Examination of the stimuli utilized in phase three indicated that there were sixteen within this range. If each of these sixteen were repeated three times, a stimuli set of forty-eight would be available. Such a set was constructed, and thirty-two males from the Stanford freshman and sophomore dormitories were recruited by advertisement. As before, each was paid $1.50 and each received a detailed explanation of the research objectives. All other procedures and conditions remained the same. The projector slide cartridge holds precisely forty-eight slides. Thus, the two demonstration stimuli used in previous tests were removed and the cartridge filled with this new set. The first two slides in the series were used as demonstration stimuli, and they were then repeated as stimuli numbers 49 and 50. A complete description of the stimuli utilized in this special set is presented on the following page. The presentation order which is noted was specifically arranged to maximize the separation between identical or closely related stimuli. In the computation of matrices, the first two stimuli are excluded (the demonstration stimuli), but they are included as the last two stimuli. The total set involved in the computations is thus composed of forty-eight stimuli.

The aggregate transition matrix obtained appears as follows.

		n + 1		
		Black	White	
n	Black	.41	.59	.45
	White	.51	.49	.55
		.46	.54	1504

$X^2 = 14.601$ $p < .001$

PHASE SEVEN STIMULI

FORM	P_w elicited in Phase Three Test	MOUNT CODE
1, first position, front mount	.56	A
1, 180 degree rotation, back mount	.47	B ✓
1, first position, back mount	.47	C ✓
1A, 180 degree rotation, front mount	.47	D ✓
1A first position, front mount	.59	E ✓
3, 180 degree rotation, back mount	.47	F ✓
3, first position, back mount	.47	G ✓
CR3, 180 degree rotation, front mount	.44	H
CR3, first position, front mount	.50	I
CR3, 180 degree rotation,back mount	.59	J
CR3, first position, back mount	.56	K ✓
CR3A, first position, back mount	.47	L ✓
CR5, 180 degree rotation, front mount	.59	M ✓
CR5, 180 degree rotation, back mount	.53	N ✓
CR5A, 180 degree rotation, front mount	.53	O ✓
CR5A, first position, back mount	.56	P

N = 16. Each stimulus to be presented three times.

PRESENTATION ORDER

1 H (dem.)	16 F	31 J	46 J
2 A (dem.)	17 K	32 E	47 L
3 P	18 M	33 O	48 C
4 F	19 H	34 G	49 Repeat 1
5 C	20 E	35 A	50 Repeat 2
6 N	21 O	36 N	
7 K	22 G	37 K	
8 B	23 B	38 D	
9 O	24 N	39 H	
10 G	25 I	40 L	
11 A	26 D	41 J	
12 M	27 P	42 E	
13 I	28 F	43 I	
14 D	29 C	44 P	
15 L	30 M	45 B	

Notice that once again there is an inflation of the probability of a white response. This occurs despite the fact that each stimulus in the set was selected from within the range .40 to .60 (proportion of white responses elicited). The only tenable conclusion is that this inflation is a property of the task as a whole, independent of the specific stimuli which compose the task. This is further substantiated by the fact that, once again, the first stimulus elicited 78% white responses. When the same stimulus was presented three additional times within the set, it elicited 63%, 41% , and 47% white responses. Many of the subjects reported that "at first, white seemed to jump out" and many reported feeling some discomfort after entertaining the hypothesis that "all were white". The white bias which appears in this matrix is, however, not sufficient to warrant dismissal of the task. The bias must be tolerated, and its real effect cannot be determined until it is employed in experimentation. At least (and this is not a small accomplishment), there is ample evidence pointing to a feature inherent in the task itself, and this knowledge can be used to more accurately evaluate the processes of experimental interest.

Documentation of this effect has now been uncovered. Tolansky (1964) has devoted a chapter of his work to "irradiation illusions". The phenomenon was studied almost a century ago by Helmholtz. Tolansky has the following to say on the problem.

> According to Helmholtz the cause of this illusion is the irradiation of light from the light regions into the dark regions. This irradiation seems to be a physical effect which takes place within the materials out of which the eye is constructed. ... The fluids in the eye are by no means of perfect clarity, and there is invariably some scattering of light on its passage from the entrance pupil to the retina at the rear of the eyeball, which is the sensitive recording device which detects the light. Then, again, the retina itself has a granular structure. The upshot seems to be that bright light appears to spill over a little into the dark regions of an image on the retina. (P. 59)

Apparently it is not going to be possible to eliminate this effect in the Spatial Judgment Task given its present form. There is a potential solution, however. A set of stimuli could be produced wherein each stimulus contains two of the rectangular black and white grids. Instead of making a choice between "black" and "white", the subject would be required to determine which grid contained "more white". Such a procedure could thereby control the irradiation effect. In addition, the task would still retain the potential of being manipulated with respect to the degree of veridicality inherent within each stimulus.[6] Under the direction of Mr. Richard Ofshe, attempts are now under way to construct such a stimuli set from those films which we now possess. For the author's current research, however, the task as developed for phase seven has been employed.

Returning to the aggregate transition matrix, it is also apparent that there is some departure from independence. Again, the dependency which exists is a minor-diagonal dependency. That is, we again find slightly too high a rate of alternation. Examination of this dependency by quarters indicates

[6] It should be mentioned that Mr. Conner has employed in his research a "near veridical" version of the present form of the Spatial Judgment Task. A report on this version is forthcoming. Mr. Gordon Lewis originally suggested the use of two grids within a single stimulus.

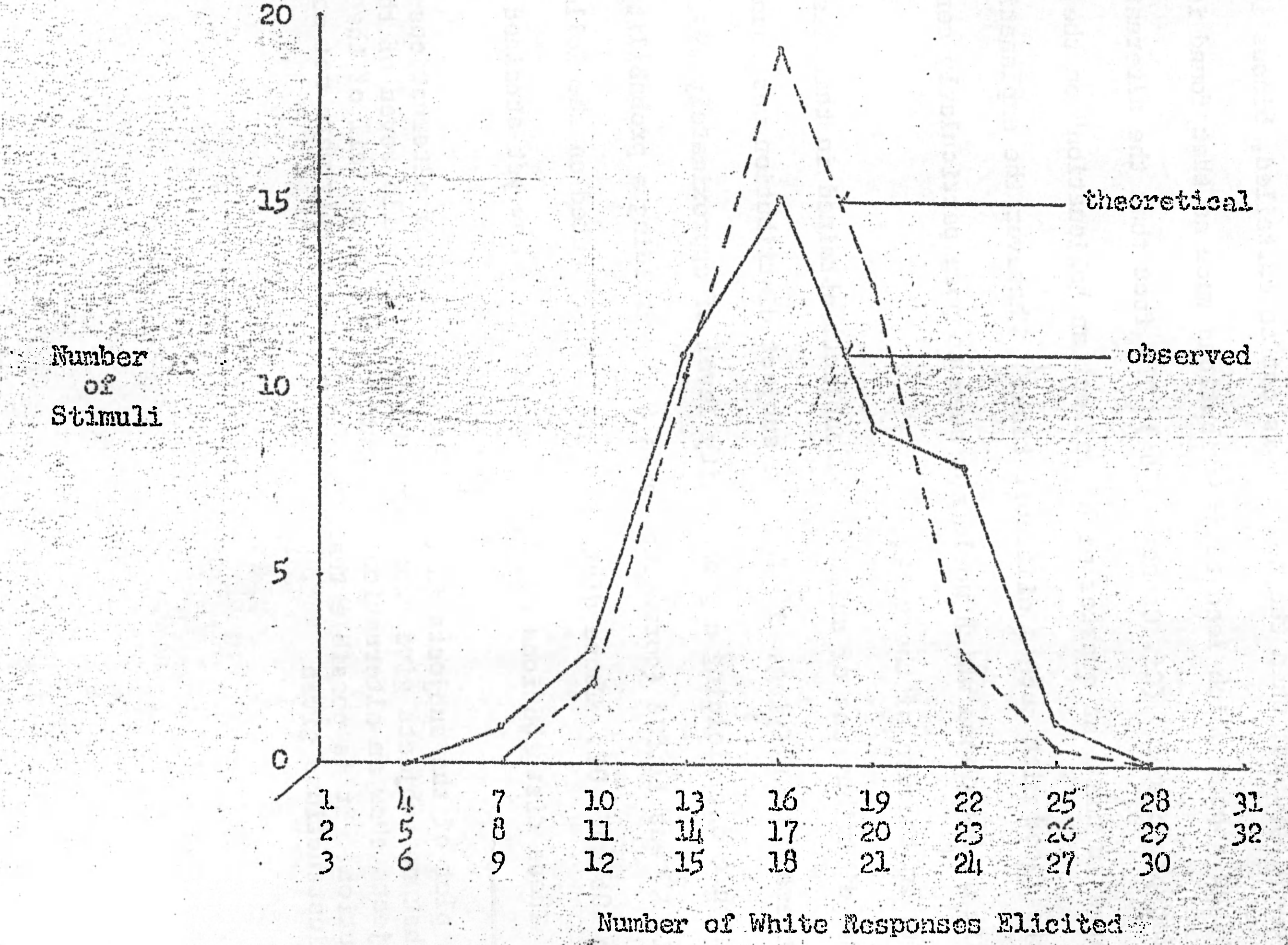
that it is present throughout the series, but increases rather sharply in the last quarter. It is difficult to account for this, but it also seems that this form of dependency is much more tolerable than a major-diagonal dependency. Previous pretest programs carried out for other tasks have indicated that this minor-diagonal dependency is a consistent phenomenon when dealing with ambiguous tasks. When the content of a stimulus is sharply curtailed, the phenomenon appears; when the content is not so curtailed, biases in the stimuli themselves appear which lead to distributions such as that found in the present case for phase six. (Cf. Conner, 1964) The fact that the alternation rate increases in the last quarter might reflect an "alienation" on the part of subjects faced with such a difficult task.[7] Whatever the explanation, this is again a phenomenon which would not seem to pose particularly serious problems in the use of the task.

The distributions of stimuli and subjects obtained in this last phase are extremely gratifying. X^2 for the subject distribution was found to be 0.875 with 4 df, yielding a probability level of approximately .90. For stimuli, X^2 was 11.506 (corrected) with 4 df, yielding a probability level between .05 and .02. These distributions are presented on the following pages. These distributions are by far the best of any yet attained for tasks

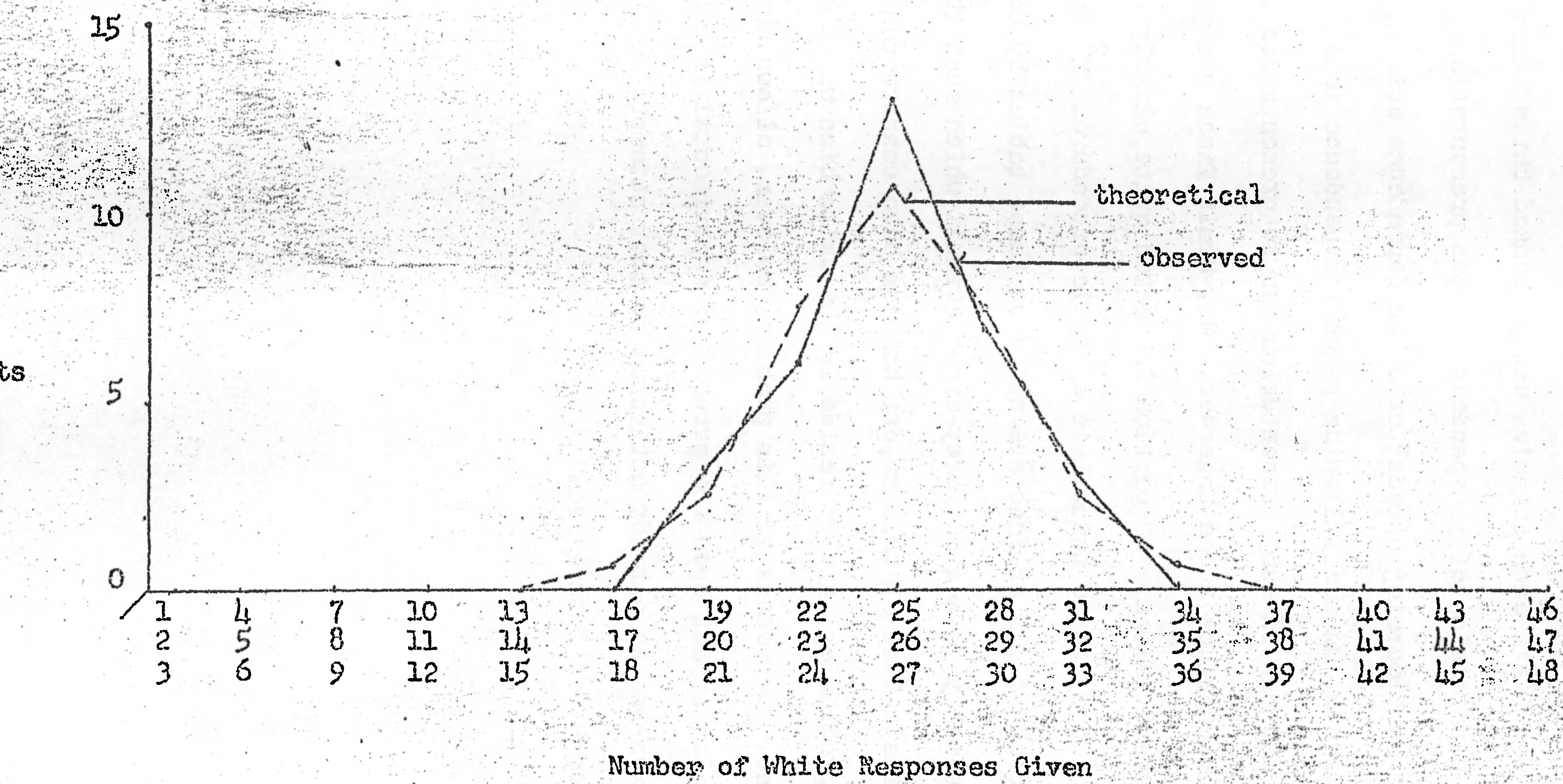
[7]A plot of the subjects in terms of the number of alternations given shows that no subjects give "too few" alternations while seven of the thirty-two subjects show an alternation rate outside the upper tail of the theoretical distribution. It is possible that these subjects simply were not motivated by the instructions given.

Distribution of Stimuli (N = 48).

Theoretical distribution is the binomial given p = .54.



Distribution of Subjects (N = 32). Theoretical distribution is the binomial given p = .54.



of this nature.

The fact that each stimulus was presented three times in this set necessitates an examination of independence across presentations. There should of course be no "learning effect" or other dependence arising from these multiple presentations. On the following pages independence is examined for presentation one to presentation two, presentation two to presentation three, and for all three presentations taken together. Mr. Hamit Fisek was gracious enough to prepare these tables. Inspection indicates quite clearly that multiple presentation of these stimuli is a viable procedure. X^2 values with a probability value of .05 or less are noted on these tables by an asterisk.

In summary, a set of forty-eight stimuli which meet the criteria specified at the outset of this report has now been constructed. The one potentially serious problem appears to reside in the white bias which was so consistently documented. Awareness of this problem, however, allows the researcher to evaluate its impact upon the process of experimental interest. In addition, this awareness has already activated further attempts to construct a more refined task for future use.

FIRST AND SECOND PRESENTATIONS

STIMULI $(P_W = 0.54)$

Response Pattern	P	Fe	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
BB	0.21	7	10	5	8	11	12	5	10	5	6	4	9	6	8	13	4	5
BW	0.25	8	6	12	6	5	9	4	8	3	3	6	4	8	9	9	4	5
WB	0.25	8	6	8	10	7	4	15	7	7	8	9	6	7	4	5	11	11
WW	0.29	9	10	7	8	9	7	8	7	17	15	13	13	11	11	5	13	11
$X^2 =$			2.396	3.015	1.253	3.535	6.140	8.807	1.854	10.932	7.267	3.687	4.848	0.711	2.711	8.169	6.347	3.585
P (df = 3)		<	.50	.50	.80	.50	.20	.05	.70	.02	.10	.30	.20	.90	.50	.05	.10	.50
		>	.30	.30	.70	.30	.10	.02	.50	.01	.05	.20	.10	.80	.30	.02	.05	.3
								*		*						*		

SECOND AND THIRD PRESENTATIONS

STIMULI ($P_W = 0.54$)

Response Pattern	P	Fe	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
BB	0.21	7	8	4	13	6	5	12	8	7	5	9	8	9	8	5	11	9
BW	0.25	8	8	9	5	12	11	8	6	5	9	4	7	4	4	13	4	7
WB	0.25	8	3	9	10	3	5	4	10	12	10	8	8	8	7	3	6	4
WW	0.29	9	13	10	4	11	11	8	8	8	8	11	9	11	13	11	11	12
X^2 =			5.044	1.646	9.544	5.711	3.265	5.682	1.253	3.236	1.307	3.015	0.267	3.015	4.044	7.265	5.229	3.646
p (df = 3) <			.20	.70	.05	.20	.50	.20	.80	.50	.80	.50	.98	.80	.30	.10	.20	.50
p (df = 3) >			.10	.50	.02	.10	.30	.10	.70	.30	.70	.30	.95	.70	.20	.05	.10	.30

FIRST, SECOND AND THIRD PRESENTATIONS

STIMULI ($P_W = 0.54$)

Response Pattern	P	Fe	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
BBB	.0973	3	5	2	6	4	4	3	7	3	1	3	4	4	5	4	2	3
BBW	.1143	4	5	3	2	7	8	2	3	2	5	1	5	2	3	9	2	2
BWB	.1143	4	2	5	4	2	4	1	5	2	2	3	2	4	5	2	2	1
BWW	.1341	4	4	7	2	3	5	3	3	1	1	3	2	4	4	7	2	4
WBB	.1143	4	3	2	7	2	1	9	1	4	4	6	4	5	3	1	9	6
WBW	.1341	4	3	6	3	5	3	6	3	3	4	3	2	2	1	4	2	5
WWB	.1341	4	1	4	6	1	1	3	5	10	8	5	6	4	2	1	4	3
WWW	.1574	5	9	3	2	8	6	5	5	7	7	8	7	7	9	4	9	8
X^2 =			8.533	5.883	10.300	9.133	9.533	11.000	8.833	14.300	10.633	6.050	5.383	3.383	8.533	14.533	13.783	7.000
p (df = 7)		<	.30	.70	.20	.30	.30	.20	.30	.05	.20	.70	.70	.90	.30	.05	.10	.50
		>	.20	.50	.10	.20	.20	.10	.20	.02	.10	.50	.50	.80	.20	.02	.05	.30
										*						*		

ADDENDUM

THE OPERATION OF THE TASK IN THE STATUS CHARACTERISTICS EXPERIMENT

The author's dissertation research employed five distinct experimental groupings. Forty-four stimuli were presented to each subject in each group. On each of the forty-four trials, subjects made an initial choice between "black" and "white". On twenty-eight of the forty-four trials, subjects in the first four conditions received a controlled disagreement from their experimental partners, and were required to make a final choice on these "critical" trials. In the fifth experimental group, however, no disagreements were exchanged. Subjects simply made an initial choice and, on the same twenty-eight critical trials, a final choice. The groups were labeled "inclusion low", "inclusion high", "burden of proof low", "burden of proof high", and "S-Alone". In this addendum, these groups will be designated one to five, respectively. The objective here is simply to evaluate the effects of the white bias inherent in the task.

Considering only <u>initial</u> choices for each of the five groups, the following aggregate transition matrices were obtained.

CONDITION 1

n \ n + 1	B	W	
B	.44	.56	.47
W	.51	.49	.53
	.48	.52	860

$X^2 = 4.621$ $.05 > p > .02$

CONDITION 2

n \ n + 1	B	W	
B	.40	.60	.40
W	.42	.58	.60
	.41	.59	946

$X^2 = 0.267$ $.70 > p > .50$

CONDITION 3

n \ n + 1	B	W	
B	.43	.57	.45
W	.49	.51	.55
	.46	.54	860

$X^2 = 3.116$ $\qquad .10 > p > .05$

CONDITION 4

n \ n + 1	B	W	
B	.41	.59	.44
W	.48	.52	.56
	.45	.55	989

$X^2 = 5.469$ $\qquad p = .02$

CONDITION 5

n \ n + 1	B	W	
B	.39	.61	.41
W	.45	.55	.59
	.43	.57	860

$X^2 = 3.575$ $\qquad .10 > p > .05$

For each group, the slightly inflated probability of a white response appears to maintain itself under the experimental conditions. For all groups combined, the probability of a white response is 0.56. If there are any departures from independence, these are not substantial enough to be of any consequence.

Having established the continuance of the white bias, a reasonable hypothesis is that subjective certainty or "confidence" in an initial choice is greater if that initial choice was "white". Data relevant to this hypothesis are presented in the following table.

		Prop. "Stay" Responses Given Black Initial Choice	Prop. "Stay" Responses Given White Initial Choice
Experimental Condition	1	.51	.67
	2	.56	.78
	3	.51	.73
	4	.59	.78
	5	.67	.82

Quite clearly there is a substantial increase in the probability of a "stay" response if the subject initially chose "white". This is true irrespective of the experimental condition. This fact, in conjunction with the fact that subjects in all conditions show the white bias, leads to the conclusion that the task had the effect of lowering the rate of change for all conditions

Finally, examination of the curves representing the probability of a "stay" response across critical trials shows that for all conditions, the probability drops sharply from trial one to trial two. For conditions two, four and five, the probability continues this downward trend from trial two to trial three. After this third critical trial, the curve stabilizes for all conditions. "First trial effects" have been consistently noted in experimental work of this nature, independent of the particular tasks employed. Nevertheless, at least part of this effect might be accounted for in terms of the present task. The fact that the curve for condition five subjects (S-Alone) demonstrates this feature as do the curves for the other four conditions suggests that this might be the case. It will be recalled that the first stimulus has consistently elicited a high proportion of white responses. Given this fact and the knowledge that the probability of a stay response is greater if the initial choice was white, it may well be that this particular task has an exceptionally great impact on initial trials. The following table presents the proportion of white responses elicited by the first four stimuli in each of the five conditions. Trials 2, 3 and 4 are "critical" trials.

	Condition				
Trial	1	2	3	4	5
1	.80	1.00	.80	.87	1.00
2	.70	.77	.80	.65	.85
3	.40	.73	.35	.48	.45
4	.50	.36	.85	.52	.50
n =	20	22	20	23	20

It appears that in all conditions, subjects tend to initially perceive the stimuli as "white". This probably leads to a high degree of confidence and a consequent low tendency to change. The extensive bias seems to dissipate quite rapidly, however, and this is consistent with the observed stabilization of the curves after the second or third critical trial. The fact that the task has some impact on initial probabilities is beyond dispute, but the extent of this impact independent of characteristic first trial effects cannot be adequately determined.

APPENDIX

Field and Laboratory Pretest Instructions

Field Phases

Instructions For Testing Spatial Judgment Ability

Excluded from Phase Two

We want to thank __________ for making (his, her) class time available so that you might participate in a study which is part of a larger study being conducted by the National Science Foundation.

Before I describe the study would you please put down your name, age, sex, major and number of (semesters, quarters) excluding this one that you have attended school at the college or university level.

I am ______ from the Sociology Department at Stanford University, and this is ______ (assistant). This (morning, afternoon) we are asking you to assist us in a study we are doing on the problem of Spatial Judgment Ability.

Let me tell you a little bit about this ability. Within the last few years social scientists have found in their studies that individuals differ in their ability to quickly and accurately perceive the spatial relationships between figures or objects. More simply, it has been found that when some individuals are presented with a set of figures or objects, they are quickly able to make accurate judgments about how those figures or objects are placed in relation to one another. Other people do not seem to have this ability to the same extent. This ability to make these accurate judgments about spatial relationships social scientists call Spatial Judgment Ability.

Frankly, at this time we do not know all the answers as to why some people have this ability more than others, although we think that it is related to background, training, and innate capacities. How important each of these things is, however, has not yet been determined.

One of the interesting things we do know is that this perceptual ability is not necessarily related to specialized skills the individual might possess, such as mathematical skills or artistic skills.

Because of the importance of this Spatial Judgment Ability, social scientists are engaged in an extensive set of studies to examine this ability among college students--such as yourselves--here and elsewhere. What we are going to do this (morning, afternoon) is to administer a specially prepared test which is used in measuring an individual's Spatial Judgment Ability.

The test consists of a series of large rectangular figures, each of which is composed of smaller rectangles either black or white in color, such as the figure you will now see. (Show slide) In each case, the smaller rectangles of one color, either black or white, are more predominant in number than the smaller rectangles of the other color. That is, one color will cover more of the area of the large rectamgle than the other color. Your task is to determine in each case which of the two colors covers the greater area. You must make your judgment rather quickly, as you will be able to examine each slide for only five seconds before it is removed from the screen. However, you should not answer immediately. Study the slide as best as you can, and at the end of five seconds I will tell you to write your answer. At that time, you will indicate your choice by circling either "black" or "white" beside the number of the slide which was presented. Please make your choice when I tell you to, not before or after. Now just for practice I want each of you to indicate your choice for this slide by circling either "black" or "white" beside number 1 on your answer sheet. (wait for approx. 5 seconds and turn off projector)

You may find that many of these slides will be difficult to judge, as the difference in the areas covered by black and white rectangles is frequently quite small. In each and every case, however, one of the colors covers a greater area. Try your best to determine which color that is, and put down an answer for every slide. Many people have told us that they felt they were simply guessing during this test, but the research indicates that even though they felt they were guessing they were probably perceiving very subtle cues which enabled them to make correct judgments. So do not worry about any uncertainty you might have, and remember to answer every slide.

Now let's quickly go through the procedure once again so that you will have everything clearly in mind. First, I will present a slide on the screen and announce its number. You will be given five seconds to study the slide. When you are instructed, you are to indicate your choice as to which color covers the greater area by circling either "black" or "white" on your answer sheet beside the number which corresponds to the number of the slide.

Before we begin the test we will go through a simple practice slide. This slide will not count on your test score; it is just for the purpose of becoming more familiar with the procedure. The next slide is slide number 2. (present)

(wait 5 seconds)

Now mark your answer sheets for slide number 2.

(hold slide for about 2 seconds and then turn projector off)

All right, in this test you will be asked to study 46 similar but different slides, and the procedure for all of them will be as I have just explained. Now is any part of the procedure unclear?...We will begin the test with slide number 3.

(present slide and wait 5 seconds)

Now, mark your answer sheets for slide number 3.

(hold slide for about 2 seconds and then turn projector off)

(wait about 5 seconds)

The next slide is number 4.

etc.

This completes the test. Will you please pass your answer sheets down the row and we will collect them.

Instructions For Laboratory Pretests of Spatial Judgment Task

I am __________ from the Sociology Department here at Stanford. This (morning, afternoon) we are asking you to assist us in a study we are doing on the problem of Spatial Judgment Ability.

IF THERE ARE TWO SUBJECTS:

[I'd like to point out that the only reason we have scheduled both of you for this hour is simply because it is more convenient for us to do so. In taking this test, you are working independently. Each of you will be making choices which have absolutely no bearing on the other person.]

I'd also like to mention just a few things about the equipment you see in front of you. This equipment is used for experiments in various types of decision-making, and is consequently more elaborate than our purposes today require. All you need be concerned with is that part of your panel which reads "Final Choice". I'll explain the use of the equipment further in a moment. Right now I want to tell you a little bit about this Spatial Judgment Ability.

Within the last few years social scientists have found in their studies that individuals differ in their ability to quickly and accurately perceive the spatial relationships between figures or objects. More simply, it has been found that when some individuals are presented with a set of figures or objects, they are quickly able to make accurate judgments about how those figures or objects are placed in relation to one another. Other people do not seem to have this ability to the same extent. This ability to make these accurate judgments about spatial relationships social scientists call Spatial Judgment Ability.

At this time we do not know all the answers as to why some people have this ability more than others, although we think that it is related to background, training, and innate capacities. How important each of these things is, however, has not yet been determined.

One of the interesting things we do know is that this perceptual ability is not necessarily related to specialized skills the individual might possess, such as mathematical skills or artistic skills.

Now what we are going to do this (morning, afternoon) is to administer a specially prepared test which is used in measuring an individual's Spatial Judgment Ability.

The test consists of a series of large rectangular figures, each of which is composed of smaller rectangles either black or white in color, such as the figure you will now see. (Show 1st demonstration slide) In each case, the smaller rectangles of one color, either black or white, are more predominant in number than the smaller rectangles of the other color. That is, one color will cover more of the area of the large rectangle than the other color. Your task is to determine in each case which of the two colors covers the greater area. You must make your judgment rather quickly, as you will be able to examine each slide for only five seconds before making a choice. However, you should not answer immediately. Study the slide as best you can, and at the end of five seconds I will tell you to make your "black" or "white" beside number 1 on your answer sheet. (wait for approx. 5 seconds and turn off projector)

You may find that many of these slides will be difficult to judge, as the difference in the areas covered by black and white rectangles is frequently quite small. In each and every case, however, one of the colors covers a greater area. Try your best to determine which color that is, and put down an answer for every slide. Many people have told us that they felt they were simply guessing during this test, but the research indicates that even though they felt they were guessing they were probably perceiving very subtle cues which enabled them to make correct judgments. So do not worry about any uncertainty you might have, and remember to answer every slide.

Now let's quickly go through the procedure once again so that you will have everything clearly in mind. First, I will present a slide on the screen and announce its number. You will be given five seconds to study the slide. When you are instructed, you are to indicate your choice as to which color covers the greater area by circling either "black" or "white" on your answer sheet beside the number which corresponds to the number of the slide.

Before we begin the test we will go through a simple practice slide. This slide will not count on your test score; it is just for the purpose of becoming more familiar with the procedure. The next slide is slide number 2.

(present)

(wait 5 seconds)

accurate judgments about spatial relationships social scientists call <u>Spatial Judgment Ability</u>.

At this time we do not know all the answers as to why some people have this ability more than others, although we think that it is related to background, training, and innate capacities. How important each of these things is, however, has not yet been determined.

One of the interesting things we do know is that this <u>perceptual</u> ability is not necessarily related to specialized skills the individual might possess, such as mathematical skills or artistic skills.

Now what we are going to do this (morning, afternoon) is to administer a specially prepared test <u>which is used</u> in measuring an individual's Spatial Judgment Ability.

The test consists of a series of large rectangular figures, each of which is composed of smaller rectangles either black or white in color, such as the figure you will now see. (Show 1st demonstration slide) In each case, the smaller rectangles of <u>one</u> color, either black or white, <u>are more predominant in number</u> than the smaller rectangles of the other color. That is, <u>one color will cover more of the area</u> of the large rectangle than the other color. Your task is to determine <u>in each case</u> which of the two colors covers the greater <u>area</u>. You must make your judgment rather quickly, as you will be able to examine each slide for only five seconds before making a choice. However, you should not answer immediately. <u>Study</u> the slide as best you can, and at the end of five seconds I will tell you to make your

Now let's quickly go through the procedure once again so that you will have everything clearly in mind. First, I will present a slide on the screen and announce its number. You will be given five seconds to study the slide. When I ask you to, you are to indicate your choice as to which color covers the greater area by pressing either the "black" or "white" button on the Final Choice portion of your panel.

Before we begin the test we will go through another practice slide. This slide will not count on your test score; it is just for the purpose of becoming more familiar with the procedure. The next slide is demonstration slide number 2. (present. wait five seconds)

Now make your choice for slide number 2. (hold slide until choices have been made; then turn projector off) [Record their choices] [Press Relay Release]

All right, in this test you will be asked to study 48 different slides, and the procedure for all of them will be as I have just explained. Now is any part of the procedure unclear? ... We will begin the test with slide number 3. (present slide and wait 5 seconds)

Now make your choice for number 3. (after choices, projector off) [Press Relay Release] (wait about 5 seconds)

The next slide is number 4.

etc.

.

.

.

(after # 48)

All right, this completes the test. I'll be happy to answer any questions you might have at this point.

REFERENCES

Conner, Thomas L. "Three Tasks for Use in Laboratory Small-Group Experiments", Laboratory for Social Research Technical Report No. 11, Stanford University, Stanford, California, December, 1964.

Goodman, Leo A. "Statistical Methods for Analyzing Processes of Change", American Journal of Sociology, 68 (July, 1962), pp. 57-78.

Tolansky, S. Optical Illusions, Pergamon Press, The Macmillian Company, New York, 1964.