THE CONTEXT EFFECT IN PROBLEM SOLVING

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY

Joseph Weldon Jennings, Jr.

1962

THE CONTEXT EFFECT IN PROBLEM SOLVING

By

Joseph Weldon Jennings, Jr.

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1962

ABSTRACT

THE CONTEXT EFFECT IN
PROBLEM SOLVING

by

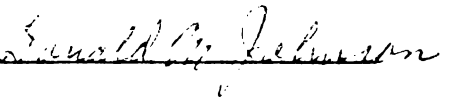
Joseph Weldon Jennings, Jr.

This thesis was concerned with the demonstration of a context effect within the problem-solving enterprise. Using the theoretical model of the problem-solving enterprise developed by D. M. Johnson, it was proposed that one of the reciprocal influences existing between the production and judgment processes in problem solving could be attributed to a context effect.

Two hypotheses were developed. The first hypothesis asserted that a self-produced stimulus context is capable of generating a context effect in the problem solver's judgment scale. The second hypothesis asserted that where a judgment process precedes a production process, the mean value of the produced stimuli will be shifted towards the mean value of the judgment context. These hypotheses were tested by comparing the performances of groups initially subjected to different stimulus contexts. All the experimental data supported the hypotheses.

Based on the findings in the present study, it was proposed that the context effect has a directional influence

on the course of a problem-solving enterprise. Essentially, with each cycle of production-judgment-production, the stimulus context means could be shifted further from the original production context. Also, investigation of the possible facilitating or hindering aspects of this directional influence was suggested.

Approved 

ACKNOWLEDGMENT

The author owes a debt of gratitude to Dr. Donald M. Johnson, the chairman of his committee. Dr. Johnson's sage counsel was vital to the planning and execution of this research, and the development of this manuscript.

Also, the author wishes to thank Dr. Charles Hanley and Dr. Stanley C. Ratner for their advice and criticism necessary for the completion of this manuscript.

****

TABLE OF CONTENTS

LIST OF TABLES

LIST OF EXAMPLES AND FIGURES

INTRODUCTION

History and Present Disposition of the Context Effect

It is the aim of this section to relate the efforts made towards achieving an understanding of the context effect. To do so it is necessary to begin with Fechner and his desire to establish an "exact science of the functional relations . . . between body and mind." Taking Weber's formula on lifted weights, $\Delta s = ks$ or $\Delta s/s = k$, Fechner derived his psychological equation $\psi = k \log s'$ (23, p. 431). Fechner held that this was the invariate relationship between the mental and the physical where ψ is the unit of mental intensity necessary for a stimulus to be judged just noticeably different from another stimulus.

Unfortunately for Fechner, his data were constantly subject to negative time-order errors. To explain this phenomenon, Fechner proposed a fading image hypothesis in 1860 (33, p. 226). The hypothesis held that on lifting the standard stimulus, the subject gained a kinesthetic "image" of that stimulus. Then, in the time between the first stimulus and the comparison stimulus, the image "faded" so that the next stimulus was overestimated. This hypothesis was agreeable to the introspectionist school until Schumann in 1898 demonstrated that "images" did not always occur (34, p. 440). Schuman's work plus that of Kulpe, Binet, and

others heralded the end of the "introspective school" but with its demise it left a host of questions that demanded answers. These questions centered around the reason for error in judgment. It was from this interest in error that an understanding of the context effect grew.

In 1899, Martin and Muller found that "absolute impressions" were involved in the judgment of weights (34, p. 440). They proposed that the subject builds a subjective scale of weights based on his experience with the stimulus series. While not directly aware of this scale, he uses it to judge any particular weight. Then in 1909, while working with the reproduction of lines, Hollingworth found that the PSE has no fixed location but lies near the middle of the range of all the stimuli used in the experiment. Also, if the stimulus range is moved up or down the PSE follows. Below the PSE or indifference point the time error is positive, while above the PSE it is negative (15, p. 323). Ipsen corroborated Hollingworth's findings and also found that the PSE could be shifted without raising or lowering the whole range. The shift could be obtained by giving the subject a disproportionate share of one high or low stimulus (34, p.446).

In 1928 Wever and Zener, using the method of single stimuli for weights, found a negative time order error. They proposed that when the subject became acquainted with the stimuli he developed a "mental formation" of an "absolute series" which was a combined trace of the whole series. With time this trace fades and results in a negative time order error (30, p. 6). Several years later Pratt (28), using a

sound pendulum, found that what Wever and Zener had called a "fading mental formation" could well be a shift in the response scale with time towards some sound commonly heard prior to the experimental situation.

Volkman's (32) use of the concept of "anchoring" in 1936 introduced an important quantitative aspect into the understanding of the context effect. He demonstrated that when one of the categories on the judgment scale was associated with the horizontal in the judgment of the inclination of lines, the remainder of the scale was extended and shifted in the direction of the anchor.

A number of studies ensued from Volkman's work but it was left to Rogers (30) to demonstrate the extent of shift in the "absolute scale" as a function of the remoteness of the anchoring stimulus. His data on judgments of the inclination of lines and the lifting of weights, using the method of single stimuli, indicated that the anchoring effect, as indicated by a shift in the category limens, is directly proportional (a rectilinear function) to the remoteness of the anchoring stimulus. Thus the "absolute scale" (see Wever and Zener) is compressed and lowered by a low anchor and raised and extended by a high anchor.

McGarvey extended these findings to verbal materials (25). By assigning an example of an occupational class or an undesirable social behavior to a category on the judgment scale, she obtained results which paralleled those reported by Rogers. McGarvey concluded that, "The extension of the absolute scale involves not only a displacement of

the scale with reference to the range of values represented by the stimulus-series, but also a widening of the categories of response . . ." (25, p.78).

Next, two major attempts to build an empirical equation for prediction of judgments from knowledge of the stimulus context becomes available. The first such attempt was presented by Johnson (15) in which he offered an equation for the prediction of category limens. He called his formulation a "generalization theory". He maintained that any stimulus X produces a "central effect" Y in the subject. It is these Y values that the subject uses to build an internal response scale. (See Martin and Muller.) The Y effect also spreads or generalizes in both directions along the apparent stimulus continuum. This generalizing spread forms a gradient symmetrical about the point of inception. The form of the gradient is determined by the response function of the receptors involved in perceiving the stimulus. As in the case of hearing, the Y function can be written, $Y=f(\log X)$. Now, where the gradients of two central effects intersect establishes the limen; i.e., the limen between a high and low pitch. This is equivalent to saying that the category limen is the log mean of all the stimuli presented for judgment (16, p. 345).

The next theory was Helson's formulation for "adaptation-level as a frame of reference" (7). Helson said this of his theory, "For every excitation-response configuration, there is assumed a stimulus which represents the pooled effect of all the stimuli and to which the organism may be

said to be attuned or adapted. Stimuli near this value fail to elicit any response from the organism or bring forth such neutral responses as indifferent, neutral, doubtful, equal, or the like, depending upon the context of stimulation." (7, p.2) Thus, the adaptation-level is a function of all the stimuli acting upon the organism at any given time, past and present.

While Johnson derived his equations on the basis of generalization gradients, Helson formulated his from the application of the Weber-Fechner law to perception of color (9). (See Fechner). Beginning with a formula derived from his work in perception, $AL = K\ (A_O\ 3\bar{A})^{\frac{1}{4}}$, Helson was able to derive the following equation for the limen in lifted weights using the method of constant stimuli: $\log (AL + 0.75d) = \left[(3\Sigma \log Xi/n) + \log c\right] /4$. C is the value of the standard stimulus, 3 is an empirical weight given to the stimulus series over the standard, /4 is the fourth root exponent carried from the arithmetical form for the proportionate weight of the stimulus series compared to the standard, and 0.75d is an empirical constant needed to reduce the AL when negative time-order errors are found in lifted weights.

This is basically the equation Johnson presented even though the two began with different theories. Johnson's equation was based on the method of single stimuli which means that there was no standard stimulus. Therefore, dropping the C term and also the constant 3 for weighting and thus the 4th root transformation term reduces the equation to, $\log (AL + 0.75d) = \Sigma \log Xi/n$. Next, Johnson was working

with pitch which involves no time-order errors. Therefore, the 0.75d constant may be dropped, leaving, log AL = log Xi/n. Essentially then, Helson's equations (separate ones are derived for each application) differ from Johnson's only in the judicious use of empirical constants and weights.

While the above formulas apply to the stimulus series to which the subject's attention is intentionally directed, Helson also identified two further sources of stimulation which have an effect upon judgment. One is the background immediately surrounding the stimuli to be judged, and the other is the residual effect of past experience with the stimuli of the same type as the subject is presently dealing with. Although Helson has attempted to derive formulas for the quantification of these two factors (8) (9) (10), under most experimental conditions they are unascertainable due to the complexity of the situation and the difficulty of assessing the previous experience of the subjects.

Helson's formula for assessing the influence of the background affords a pooled estimate. Engel and Parducci (2), using a novel stimulus situation in which the background could be identified somewhat more readily than is usually the case, conclude that Helson's theory must be modified so that the value of each stimulus in the equation would be defined with respect to a weighted mean of the stimulus and its simultaneously present background.

Also, both Johnson and Helson considered the geometric mean the primary measure of the context effect. But Parducci and Marshall (26) (27) report that the residual

error found in the comparison between AL predictions and observed results in psychophysical data may be due to the fact that the predictive value of the geometric mean is dependent upon the mean's relationship to the median of the stimulus context and the midpoint of the two end stimuli. Apparently, the mean of the midpoint and median gives a better approximation to the true AL. This is due to two conflicting tendencies: (1) to divide the range into proportionate subranges and (2) to use the alternative categories with proportionate frequency. These two tendencies tend to balance one another so that the mean of their two indicators, the midpoint and the median, gives the best approximation to the AL.

It is apparent that the equations of Johnson and Helson have limited applicability for most research purposes which do not deal with psychophysical judgments under extremely well understood stimulus conditions and involving known receptor functions. Therefore, most research on the context effect involves the measurement of the differential effect of separate or partially overlapping segments of the stimulus continuum upon the judgment of stimuli common to or included in the two stimulus series.

Fehrer (3) has demonstrated the context effect in the rating of attitude statements. Using Thurstone's "Attitudes towards War" scales she was able to show that when originally moderate items were rated in the context of highly militaristic items they were rated less militaristic than when rated in the context of a series of pacifistic

items. But, she also found that extreme items of either type when rated in contrary contexts were not displaced. Fehrer held that these exceptions were due to the judges having re-defined the extreme categories in some manner that exempted them from being influenced by the remainder of the scale.

Campbell, Lewis and Hunt (1) addressed themselves to the problem such as that encountered by Fehrer. They suggested the possibility that the context effect might be an artifact because the subject is regularly forced to use a very limited number of categories in which to classify the stimuli and a language for identifying these categories which is novel to them and relevant only in the experimental situation. They proposed that, if this were the case, supplying the subject with a great number of categories and a language which was "absolute, extensive, and extra- experimentally anchored" should eliminate the context effect. They supplied their subjects with a cardboard copy of a piano key board on which the subjects indicated the position of the key which would correspond to a electronically produced tone. By slowly descending or ascending the scale in random patterns of tones they were able to demonstrate a shift in categories assigned to middle-range tones towards the upper or lower end of the key board respectively. Thus, they concluded that the context effect was not an artifact of the experimental procedure.

A study by Fine and Haggard (4) lends support to the Campbell, Lewis, and Hunt experiment. The former two workers

used various types of foods as stimuli such as roast beef, fried liver, stewed kidneys. Because people often discuss the relative merits of various common foods, the task was not entirely unique and relative only to the experimental situation. Fine and Haggard's hypothesis that the scale value of adjectives rated in the context of foods would increase signficantly in a more specific context of a highly acceptable food and decrease when rated in the context of an unacceptable food was supported by the data.

The context effect as an explanatory principle has found acceptance in many other areas of behavioral research, especially in social psychology (11) (12) (24) (31). In a later section, the argument will be developed that the context effect is a factor to be considered in problem solving as one of the manifold influences the problem solver is necessarily subjected to in attempting to secure a solution.

A Model of the Problem-Solving Enterprise

Johnson has developed a tripartite model of the problem solving enterprise (16) (17) (21). The three parts or processes of this model are preparation, production, and judgment. During the preparation process the problem takes shape for the problem solver. Essentially, on recognition by the problem solver that habitual responses are in some manner ineffectual in a new situation, he begins structuring or organizing the relevant materials in terms of the instructions given him, as in an experiment, or in terms of his motives. (16) In so doing a "task-attitude" or "set" is

adopted which serves to limit the effective environment which must be taken into account. The set determines, to a great extent, the type of responses the problem solver will make and the manner in which further preparation will be accomplished. (6) (18) (19).

A further result of the task-set established during the preparation phase is the development of a search model which is used much as one would use a partially constructed jigsaw puzzle against which the remaining pieces are matched. The activity directed toward the development of ideas to fit the search model is appropriately called the "production process" (17). During this productive phase the problem solver conceives of various alternatives which might serve as solutions. Under the aegis of the search model, production will be limited to conform to his understanding of the problem. The restriction placed on the production process results in ideas or stimuli which have some relationship to one another such as machine tools or names for a male heir.

The third process is judgment. This is a "conclusive or decisive process" (16, p. 282) by which ideas or stimuli are assigned to various categories along the dimension which underlies the production process. The judgment process is a necessary terminus for a problem in that the problem solver cannot implement the manifold possibilities developed during the production process. Therefore, he must embark upon that course of action which he judges most likely to end in success. But the first judgment need not be the last for, "Judgment may also occur at any point in a complicated

problem-solving enterprise, as when one has to decide between two promising initial strategies; hence judgment may precede production as well as follow it . . ." (21, p. 129). Thus it can be said that in fairly complex problems, production and judgment may alternate or follow a reciprocal pattern.

Of necessity, the model gives a simplified picture of the total enterprise. In actuality, considerable temporal overlap of the three processes generally exists. But, Johnson has devised an experimental method of serially analysing the problem-solving enterprise so that the overlap amongst the three phases is considerably diminished. This is accomplished by withholding information and/or restricting activity necessary to one process until the preceding one has been accomplished (17) (20) (22). Factor analytic techniques applied to the results of one such serial analysis strengthens the model in that three major factors resulted which were clearly identifiable as preparation, production, and judgment (22).

Essentially, Johnson's model and experimental technique is an attempt, "to describe problem-solving activities in functional units that are larger than the single response and smaller than the whole problem-solving episode" (17, p. 66). This approach finds its antecedents in the work of John Dewey and Graham Wallas. In its express aim of illuminating the psychology of problem solving, this approach is superior to techniques which treat the problem solving enterprise as an undifferentiated whole and use over-all measures of outcome

as seen in the work of Rimoldi (29) and John (13) (14). These latter approaches tend to become techniques for assessing individual differences in problem-solving ability and do not help identify and explain the nature of the various component processes which taken together may be called a "problem-solving episode".

The Context Effect in the Problem-Solving Enterprise

From Johnson's model of the problem-solving enterprise discussed in the preceding section, the production process has been identified as a source of stimuli which serve as possible solutions to the problem. Because these produced stimuli are related in a way dictated by the set and search model, they have some common dimension which maintains their relation to the problem at hand. Thus, the produced stimuli bear the essential properties of a stimulus series such as encountered in the first section in the discussion of the context effect. What is novel about this produced stimulus series with respect to the stimulus series commonly presented to the subject in the investigation of the context effect is that the problem solver is the source of the stimulus series or context upon which he makes his judgment. If Johnson is correct in his analysis of the production process there is theoretically no reason for assuming that the novel source of the stimulus context developed in that process should preclude a "central" or context effect from being established. This effect should be expressed in a predictable shift in a subject's categorization of produced stimuli.

To give an example of what is meant, consider the situation where the problem solver produces only poor possibilities, of which a disproportionate number are very poor indeed, based upon some absolute scale which the problem solver necessarily does not know of or he would not have a problem. From Ipsen's work, for example, it is known that this will result in the PSE or category limen being shifted downward. This increases the possibility of one of the less poor items being judged acceptable or worth attempting. This assumption will be tested experimentally.

Next, with reference to Johnson's model of the complex problem-solving episode or enterprise, it was stated that production and judgment reciprocate such that a judgment process can precede another production process. In this situation, the judgment process establishes a context effect as previously discussed. Thus, this context effect becomes an antecedent condition under which the subsequent production process operates. Because the judgment process is also a decisive process, this should affect the ensuing production process, for the decision made during the judgment process determines what sort of ideas or stimuli may better serve as a solution. But the judgment process is assumed to be under the control of the stimulus context afforded for judgment and, if this is the case, the ensuing stimuli or ideas which will be produced should then be under the influence of the stimulus context that existed during the preceding judgment process. Essentially, it is expected that the mean value of the ideas or stimuli produced during the ensuing

production phase should be nearer the mean value of the stimuli of the preceding judgment scale if the context effect is transferable from the judgment process to the production process, than if it is not transferable.

If the reciprocal influence of the production and judgment processes upon one another can be demonstrated experimentally as being a function of the context effect, this will serve to identify the context effect as a systematic influence operating within the problem-solving enterprise. Experimental demonstration of this systematic influence will open the way for further appraisal of the continuity and directional aspects of the problem-solving episode.

Hypotheses

Based on the preceding discussion of the probable influence of the context effect upon the problem-solving enterprise as it is conceived of in the model developed by Johnson, two hypotheses were developed:

(1) When production of solutions to a problem is followed by judgment of solutions, the solutions produced constitute the context within which the solutions are judged. Hence the category limen of the scale of judgment will shift towards the mean scale value of the solutions produced.

(2) When judgment precedes production, the objects judged constitute a context which influences production. Hence the mean scale value of the objects produced will shift toward the mean of the objects judged.

The experiment designed to test these hypotheses uses a procedure whereby the predicted changes in behavior are

a function of contrasting stimulus contexts. This experiment uses non-verbal stimuli. An additional experiment, which will also be reported, was designed to extend the generality of the findings from the test of the second hypothesis into the realm of verbal materials.

EXPERIMENT I

Design

For this experiment, four groups of ten subjects each were used. The subjects were research volunteers from the Introductory Psychology course. Subjects were randomly distributed into one of the four groups. All subjects were run on an individual basis.

To test the first hypothesis, two groups of subjects were required to produce a series of angles and then judge a series of angles. The two groups of subjects produced different ranges of angle sizes; one group produced angles which were acute, the other group produced angles which were obtuse. Thus, these two groups produced different stimulus contexts. Then, in order to test the hypothesis that the self-produced stimulus context results in a context effect, which will shift the category limen of the subject's internal judgments scale towards the mean value of the self-produced stimulus context, both groups were asked to judge a common stimulus series of angles having a range which spanned the ranges of the two self-produced stimulus series. If the hypothesis is correct, the group which produced the obtuse angle context should evince a category limen higher than the group which produced the acute stimulus context.

The remaining two groups of subjects were used in the test of the second hypothesis. Each group of subjects judged a different range of stimuli. One group judged the angles of an obtuse series either "large" or "small". The other group judged the angles of an acute series either "large" or "small". Thus, the two stimulus series constitute differing contexts for judgment, although not self-produced as would be the case in solving a problem, and the resulting average category limens of the two groups will differ. The subjects were then called upon to produce angles with no specification as to size being made by the experimenter. If the hypothesis is correct, the differing stimulus contexts of the proceding judgments will affect the production process so that the mean angle produced by the group which judged obtuse angles "large" or "small" will be larger than the mean angle produced by the group which judged the acute angle series.

Material

For use in testing the first hypothesis, one series of angles was constructed, acute through obtuse. This Standard or common series ranged in size from 20^{o} to 160^{o} in 20^{o} steps which afforded a mean of 90^{o} for the series. Each of the 8 angles comprising the series was drawn on 3" X 5" blank, white cards using black india ink. The arms of the angles were 2 inches in length and approximately 1 mm. wide. The apexes of the angles in this series were varied right, left, up, and down. The reason for this will be explained

in the Procedure section.

For use in testing the second hypothesis, two series of 12 angles were prepared. One series consisted of acute angles having a range from 5^{o} through 55^{o} with a mean of 30^{o}. The other series consisted of obtuse angles having a range from 125^{o} through 175^{o} with a mean of 150^{o}. These angles were constructed in the same manner as the angles used in testing the first hypothesis.

Procedure

As noted in the Design section, four groups of subjects were used in this experiment. Two of these four groups of subjects were required to produce a series of twelve angles, then judge a series of 16 angles, thus comprising that part of the experiment aimed at testing the first hypothesis. The subjects in one group were persuaded to produce a series of twelve obtuse angles ranging in size from about 125^{o} to 175^{o}. Production consisted of the freehand drawing of angles. The production was controlled by experimenter comments such as, "not so small", "a little larger than the last one", etc. No problem was encountered in having the subjects draw twelve such angles. Each angle was drawn on a separate sheet of a blank, white 5" X 8" tablet. This group was labeled the PJ-o group (signifying Production followed by Judgment in an obtuse context).

The subjects in the other group were also called upon to produce a series of twelve angles, but this group was persuaded by E's comments to draw acute angles ranging in size from about 5^{o} to 55^{o}. This group constituted the PJ-a

group (signifying Production followed by Judgment in an acute context).

After having encouraged the PJ-a and the PJ-o groups to produce different series of angles, the two groups were treated in an identical manner. Both were asked to respond by saying "large" or "small" to the "acute through obtuse" series of 8 angles. The angles were presented one at a time in random order with the subject responding to each angle at the time of presentation. The series of angles was then repeated with the subject again responding "large" or "small" to each angle. Interpretation of "large" and "small" was left entirely to the subject.

As was noted in the Materials section, the apexes of the angles in this series were varied right, left, up, and down. This varying of the direction of the apexes was done to "orient" the subjects toward some uniform rationale for this phase of the experiment which, while incidental to the experiment, would distract the subject from formulating some other rationales bizarre enough to distort the subject's perception of the task in some unknown manner. After having judged the series of angles and thus completing the experiment, each subject was asked what he thought the purpose of the experiment was. Those who had any thoughts on the matter at all invariably reported that the experiment seemed to center around the direction of the apex.

The two remaining groups judged series of angles and then produced angles, and constituted that part of the experiment aimed at testing the second hypothesis. These two

groups received different series of angles for judgment. One group was shown the "obtuse angle series" and therefore this group was labeled JP-o (signifying Judgment followed by Production in an obtuse context.) The other group of subjects were shown the "acute angle series" and this group was labeled JP-a (signifying Judgment followed by Production in an acute context.)

The appropriate set of angles were presented to a subject in groups JP-o and JP-a. The subject had been previously instructed to judge each angle "large" or "small" at the time of its presentation. The interpretation of "large" and "small" was left entirely to the subject.

Having done this, the subject was then called upon to produce a series of twelve angles by drawing them freehand on a 5" X 8" blank, white tablet. Each angle produced was drawn on an individual sheet of the tablet. Before each angle was produced, the experimenter instructed the subject to draw the apex of the angle in one of four directions, i.e., right, left, up, or down. The direction of the apex was randomly determined. This technique of varying the direction of the apexes of the angles was successful in establishing an incidental orientation, as it was with the PJ subjects, as evidenced by the reports of those subjects who had formulated a rationale for this last phase.

EXPERIMENT II

Design

This experiment was designed solely to extend the generality of the findings from the test of the second hypothesis into the realm of verbal materials. The two groups used in this experiment were treated under the judgment-production paradigm followed in Experiment I where non-verbal material was used. The two groups received different stimulus series for judgment and then produced stimuli of the same type as presented for judgment. One group of 31 subjects received a highly offensive series of statements for judgment and were labeled JP-h while the other group of 27 subjects received a mildly offensive series and were labeled JP-m. While the subjects in Experiment I were run on an individual basis, this experiment was accomplished on a group basis using entire classes from the Introductory course in Psychology.

Material

Examples I and II are samples of the two forms of the experimental materials used. The difference between these two forms lies in the series of offensive statements on the front of each sheet. These two series were constructed from a list of 187 offensive statements compiled by McGarvey (25). The fifteen items comprising each form were selected from McGarvey's list on the basis of ratings assigned to them by a standardization group.

The standardization group consisted of ten undergraduate students. They were supplied with a copy of McGarvey's entire list of statements, and a sheet on which were instructions to use a scale of 0-100 and rate mild offenses from 10 to 30, moderately serious offenses from 40-60, and very serious offenses from 70-90. Also, the sheet afforded them space to write the rating of each statement beside the number of the statement as it appeared in McGarvey's list (see Example III). From the standardization group's ratings the median rating of each statement was calculated.

With this information, two lists of 15 statements each were compiled. One list, Example I, was constructed from highly offensive items so that the mean of the items was 75. The other list, Example II, was constructed from mildly offensive items so that the mean of these items was 36. It was hoped that the later list might be constructed so as to have a mean of 25, but this was not possible due to the paucity of items with ratings in the lower end of the scale.

Both the highly offensive and mildly offensive lists were preceded by the following instructions:

> Do This Side First
>
> Below is a list of phrases describing offensive acts. Assume the role of an observer of human behavior and rate these phrases so as to indicate how offensive most people would consider them.
>
> Please rate the offenses on a scale of 0-100. Mild offenses should be rated, say 10 to 30, moderately serious offenses 40 to 60, very serious offenses 70 to 90. Mark your rating in the space in front of each item.

Inspection of the reverse side of Examples I and II

reveals the instructions and space provided for the production phase. The instructions read, "we shall need a larger number of phrases to continue our research. Will you write, below, 12 short phrases, each describing a different offensive act." The instructions are followed by the numbers 1 through 12 in order down the left side of the sheet with space beside each number sufficient for writing a phrase. This was placed on the back of the sheet so as to diminish to a minimum the time required for the subject to proceed from judgment to production.

Procedure

As all the pertinent instructions were printed on the material the subject was to receive, (see Example I and II) the experimenter simply introduced himself and explained in general terms that the subjects were being asked to cooperate in a psychological experiment, and stressed the point that the instruction "Do this side first" should be followed. The experimenter then distributed the experimental materials to the class by alternating the rows which received the highly offensive series with the rows receiving the mildly offensive series. The subjects were not aware of this alternation in forms as the experimenter distributed one form to alternate rows at the same time as an assistant distributed the other form to the remaining rows. In this way it appeared to the subjects that the alternation in rows by the experimenter and his assistant was only a means of distributing the material in a minimal amount of time. Also,

by giving the same experimental material to each row, all subjects to the right or left of any particular subject were in possession of the same material as a casual glance might reveal, although no conversation between subjects was permitted and none observed.

As soon as the subjects received the material, they commenced judgment of the fifteen statements on the front and, on completing that, turned the sheet over and began the production of offensive statements. Production consisted of writing offensive statements. As each subject finished the experiment, he immediately turned in all material to the experimenter and left the room. The subjects took from 15 to 25 minutes to complete the experiment.

Example I. Copy of the Experimental Material containing Highly Offensive Statements used in Experiment II

[illegible]

[illegible]

Please rate the offenses on a scale of 0 - 100. [illegible]

1. [illegible]

2. [illegible]

3. [illegible]

4. [illegible]

5. [illegible]

6. [illegible]

7. [illegible]

8. [illegible]

9. [illegible]

10. [illegible]

11. [illegible]

12. [illegible]

13. [illegible]

14. [illegible]

15. [illegible]

Example II. Copy of the Experimental Material containing Mildly Offensive Statements used in Experiment II

Example III. Copy of the Sheet used by the Standardization Group to record their ratings of Offensive Statements

Experiment: Offense Rating Subject # _____; M ____, F ____

Please rate the offenses on a scale of 0 - 100. Mild offenses should be rated, say, 10 to 30, moderately serious offenses 40 to 60, very serious offenses 70 to 90, and so on.

Item No.	Your rating	Item No.	Your rating	Item No.	Your rating	Item No.	Your rating	Item No.	Your rating	Item No.	Your rating	Item No.	Your rating
1.		29.		57.		85.		113.		141.		169.	
2.		30.		58.		86.		114.		142.		170.	
3.		31.		59.		87.		115.		143.		171.	
4.		32.		60.		88.		116.		144.		172.	
5.		33.		61.		89.		117.		145.		173.	
6.		34.		62.		90.		118.		146.		174.	
7.		35.		63.		91.		119.		147.		175.	
8.		36.		64.		92.		120.		148.		176.	
9.		37.		65.		93.		121.		149.		177.	
10.		38.		66.		94.		122.		150.		178.	
11.		39.		67.		95.		123.		151.		179.	
12.		40.		68.		96.		124.		152.		180.	
13.		41.		69.		97.		125.		153.		181.	
14.		42.		70.		98.		126.		154.		182.	
15.		43.		71.		99.		127.		155.		183.	
16.		44.		72.		100.		128.		156.		184.	
17.		45.		73.		101.		129.		157.		185.	
18.		46.		74.		102.		130.		158.		186.	
19.		47.		75.		103.		131.		159.		187.	
20.		48.		76.		104.		132.		160.		188.	
21.		49.		77.		105.		133.		161.		189.	
22.		50.		78.		106.		134.		162.		190.	
23.		51.		79.		107.		135.		163.		191.	
24.		52.		80.		108.		136.		164.		192.	
25.		53.		81.		109.		137.		165.		193.	
26.		54.		82.		110.		138.		166.		194.	
27.		55.		83.		111.		139.		167.		195.	
28.		56.		84.		112.		140.		168.		196.	

RESULTS

EXPERIMENT I

Results from the test of the first hypothesis are presented in Table I. The dependent variable chosen for analysis was the number of angles judged "large" by each subject in each experimental group, PJ-o and PJ-a. This measure was chosen as the most direct, because the results required no further transformation or manipulation before the application of statistical tests. Further, a difference between the mean frequency of "large" judgments by the two experimental groups could only occur if the average category limen of the PJ-o subjects was higher than the average category limen of the PJ-a subjects. The group limens in the form of a graph which also includes the production context means. In Figure I, this graph is labeled "Production followed by judgment."

TABLE I

ANALYSIS OF THE RESULTS FROM THE TEST OF THE FIRST HYPOTHESIS

	Group	
	PJ-o	PJ-a
Group Limen	89.0^{o}	69.0^{o}
Mean	8.1	10.1
S.D.	2.08	1.45
N	10	10

$t = -2.50$; $P < .05$

Inspection of Table I reveals that the results are in agreement with the predictions. The average subject in the PJ-o group judged fewer angles "large" than did the average subject in the PJ-a group. This indicates the category limen of the PJ-o subjects was higher than that of the PJ-a subjects.

Results from the test of the second hypothesis are presented in Table II. In this situation, the measure of the dependent variable was the average angle produced by each subject in the two experimental groups. To reiterate, the expectation was that the average angle produced by the subjects in the JP-o group will be larger than the average angle produced by the subjects in the JP-a group. Such an effect is attributable to the difference in stimulus contexts experienced by the two groups of subjects during the judgment phase.

TABLE II

ANALYSIS OF RESULTS FROM THE TEST OF THE SECOND HYPOTHESIS

	Group	
	JP-o	JP-a
Mean	83.24^{o}	42.72^{o}
S.D.	18.24^{o}	10.06^{o}
N	10	10

$t = 6.15$; $P < .0005$

Analysis of the data substantiates the hypothesis. Because the JP-o subjects produced larger angles than the JP-a subjects it can be said that the context effect established during judgment is transferred to the production process. In Figure I, the graph labeled "Judgment followed by Production" displays the judgment context means of groups JP-o and JP-a and the two groups' resultant production context means.

EXPERIMENT II

It will be recalled that this experiment consists of a further test of the second hypothesis using verbal material instead of angles. In this experiment, the subjects' behavior could not be directly measured. Instead, each statement produced by the subjects was rated by the experimenter and Dominic J. Zerbolio, also a graduate student, on the same scale as used by the subjects during the first phase of this experiment. To control possible biasing or "halo" effects on

the part of the raters all of the subjects' statements were rated "blind". For this purpose, a typist copied each statement onto a separate 4 X 6 inch piece of paper and marked each sheet with a code number indicating the experimental group and subject from which the statement came. This code was devised by D.M. Johnson and its key was not revealed to the experimenter until after all ratings were completed. On completion of typing and coding, all the pieces of paper were thoroughly shuffled. The interrater reliability based on product-moment correlation between the mean of the ratings given each subject's production by the two raters was + .79, the probability of a chance occurence being less than .0005. With so high an interrater reliability, the average of the two raters' ratings of each item became the value assigned to the item.

The basic observational unit used in the statistical analysis of the difference between JP-h and JP-m groups was the mean of the ratings given each subject's statements. The subjects' means were then used in the calculation of their respective group means. Table III gives the pertinent statistics on the comparison of the two group means.

TABLE III

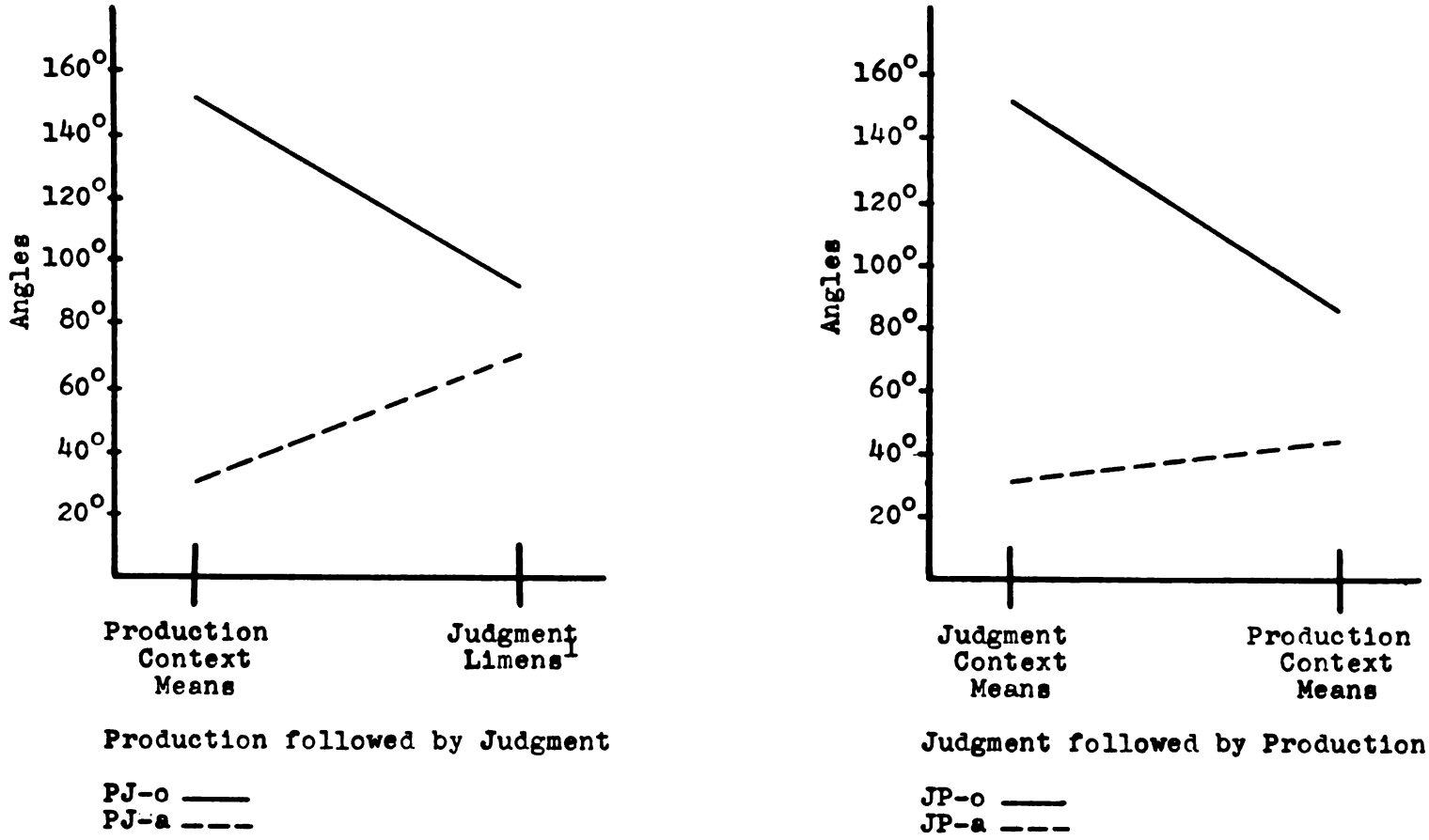
ANALYSIS OF RESULTS FROM THE TEST OF THE SECOND HYPOTHESIS USING VERBAL MATERIAL

	Group	
	JP-h	JP-m
Mean	54.29	48.02
S.D.	9.49	8.87
N	31	27

$t = 2.59$; $P < .01$

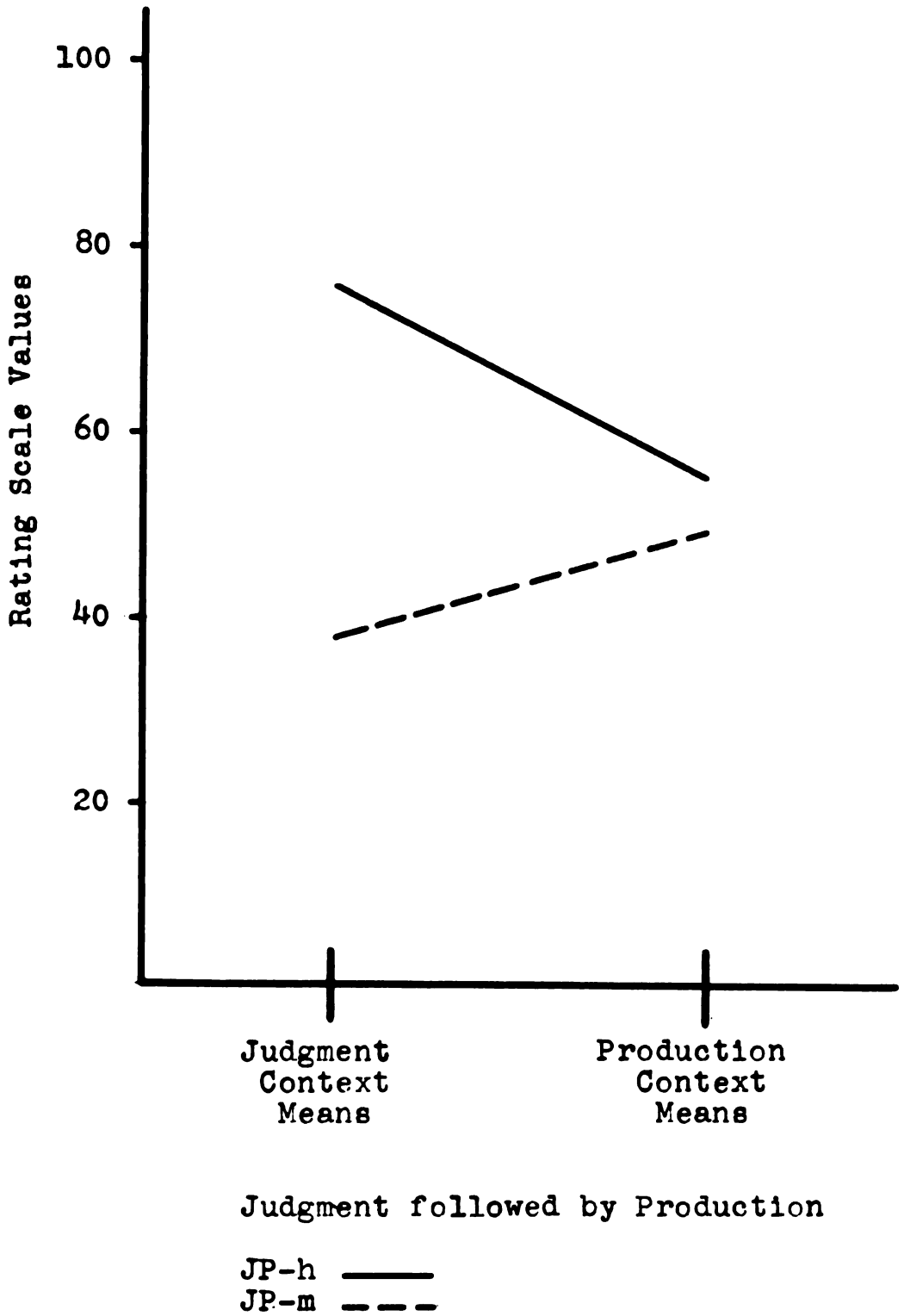
Again, the second hypothesis is supported by the data. Thus, with verbal material as well as non-verbal material, it may be concluded that the context effect established during judgment is transferred to the production process.

Figure 1. The Context Effect in Experiment I



1. Limens computed by the Spearman distribution method (34,p.207)

Figure 2. The Context Effect in Experiment II



DISCUSSION

The purpose of this thesis was to elaborate upon Johnson's model of the problem-solving episode (16) (17). The fundamental theoretical validity of this model was accepted for the present. A good deal of valuable research on problem solving has originated from the theoretical framework supplied by Johnson's model (18) (19) (22).

The focus of the present research was on the possible existence of a systematic influence operating between the production and judgment processes of the problem-solving enterprise. In his quest for a solution, the problem solver goes through one or more sequences of producing and evaluating ideas, cues, or stimuli. If this is at all a fitting description of the situation, it seemed reasonable to assume that the quality of each idea could affect the evaluation and production of other ideas. The assumption did not spring full-fledged from the experimenter's head like a Pallas Athene. The precedents for this line of thought lay in the many and varied studies of the context effect in judgment (2) (3) (4) (12) (15). The parallel is easily proposed between the situation where the experimenter supplies the stimulus context for judgment and one where the subject himself produces the stimulus context as in problem solving. But the test of the proposed parallel depends upon

the demonstration that the self-produced stimulus context will result in a context effect in the subject's judgment scale.

This proposal was embodied in the first hypothesis which stated, in essence, that a subject-produced stimulus context results in a shift in the subject's judgment limen towards the mean of the produced context. The data of Table I supports this hypothesis. Therefore, with reference to Johnson's model of the problem-solving enterprise, it seems justifiable to conclude that in solving a problem, the ideas, cue, or stimuli which the problem solver produces influences his evaluation of these ideas, cues, or stimuli during the judgment process. This influence acting upon the problem solver's judgment scale can be called a context effect.

As previously discussed, Johnson has proposed that in complex problems the production of ideas or stimuli may begin again after an initial sequence of production followed by judgment (21). This repetition of the production process represents an attempt by the problem solver to further exhaust the possibilities for additional ideas, cues, or stimuli for use as solutions to the problem. But, because the judgment process which preceded this new production process was an evaluative or decisive process the problem solver has made a tentative decision as to what may be likely solutions to the problem. The tentativeness of this interim judgment is evidenced by the fact that production of ideas or stimuli is resumed.

Based on the results obtained from the test of the first hypothesis, the evaluations made in the interim judgment process are under the influence of the stimulus context produced during the production process which preceded it. Following this line of reasoning, it was proposed that when a judgment process precedes a production process the stimulus context which existed during the judgment process may have an effect upon the ensuing production process such that the mean value of the stimuli subsequently produced will be shifted towards the mean value of the judgment context. This proposal was embodied in the second hypothesis. The data of Tables II and III support this hypothesis. Because the second hypothesis was somewhat more novel than the first hypothesis, the second hypothesis was tested using both non-verbal and verbal stimuli.

The inferences developed from the test of the second hypothesis apply only to the situation where the problem solver's set does not change from the judgment process to the production process. The second hypothesis describes a situation in which the problem solver continues the production of the same type of stimuli so that a common stimulus dimension exists between the stimuli constituting the judgment context and the stimuli resulting from the ensuing production process. When a change in set occurs, "the activity changes in kind, as when one who has just produced a block of bird names now produces a block of mammal names . . ." (16, p. 195). The possibility of a context effect persisting through a change in set is not discarded. But, the present research

was not designed to cope with the radical behavior changes encountered in such a situation.

In general, the present research would appear to justify the proposal that the context effect has a directional influence on the general course of the problem solving enterprise. A production-process stimulus context produces a shift in the limen of the ensuing judgment process. The stimulus context of a judgment process produces a shift in the mean value of the ensuing production process. Each of these shifts is in the direction of the mean of the stimulus context of the preceding process. Thus, with each cycle of production-judgment-production, the stimulus context means could be shifted further from the original production context mean. In this manner a trend or direction may be established in the problem solver's quest for a solution. Whether the directional aspect of the context effect has a beneficial, detrimental or neutral influence on the problem solving enterprise is a matter for further research.

SUMMARY AND CONCLUSIONS

Previous research on the context effect in judgment appeared to hold some significance for the further understanding of the problem-solving enterprise. Using Johnson's theoretical model of the problem-solving enterprise, the proposition was developed that one of the reciprocal influences existing between the production and judgment processes in problem solving could be attributed to a context effect.

Two hypotheses were developed. The first hypothesis was essentially a proposal that a self-produced stimulus context is capable of generating a context effect in the problem solver's judgment scale. The hypothesis found experimental support in a test using non-verbal material. The second hypothesis asserted that where a judgment process precedes a production process the mean value of the produced stimuli will be shifted towards the mean value of the judgment context. This hypothesis was tested using both verbal and non-verbal material. The experimental data supported the hypothesis.

In discussing the implications of the present research, the conclusion was reached that the context effect has a directional influence upon the course of a problem solving enterprise. Each of the above described "shifts" changes

succeeding context means away from the initial self-produced context mean. A sufficient number of such "shifts" could constitute an identifiable trend or direction in the succession of ideas produced by the problem solver. Whether this directional influence serves to expedite or hinder the problem-solving process is a question requiring further research.

BIBLIOGRAPHY

1. Campbell, D. T., Lewis, N. A. and Hunt, W. A. Context effects with judgmental language that is absolute, extensive, and extra-experimentally anchored. J. exp. Psychol., 1958, 55, 220-228.

2. Engel, Gloria and Parducci, A. Value of background in the specification of the stimulus for judgment. Amer. J. Psychol., 1961, 74, 569-575.

3. Fehrer, Elizabeth. Shifts in scale values of attitude statements as a function of the composition of the scale. J. exp. Psychol., 1952, 44, 179-188.

4. Fine, B. J. and Haggard, D. F. Contextual effects in scaling, J. applied Psychol., 1958, 42, 247-251.

5. Guilford, J. P. Psychometric Methods, New York, McGraw-Hill, 1954.

6. Hall, E. R. Formulation and reformulation of verbal concepts. Unpublished M.A. Thesis, Mich. St. Univ., 1960.

7. Helson, H. Adaptation-level as a frame of reference for prediction of psychophysical data. Amer. J. Psychol., 1947, 60, 1-29.

8. Helson, H. Adaptation-level as a basis for a quantitative theory of frames of reference. Psychol. Rev., 1948, 55, 297-313.

9. Helson, H. Adaptation level theory. in Psychology, a Study of a Science, vol. 1, S. Koch, (ed.), New York, McGraw-Hill, 1959.

10. Helson,H., Blake, R. R., Mouton, J. S. and Olmstead, J. Attitudes as adjustments to stimulus, background, and residual factors, J. abnor. and soc. Psychol., 1956, 52, 314-322.

11. Helson, H., Blake, R. R. and Mouton, J. S. An experimental investigation of the effectiveness of the "Big Lie" in shifting attitudes. J. soc. Psychol., 1958, 48, 51-60.

12. Helson, H., Blake, R. R. and Mouton, J. S. Petition-signing as adjustment to situational and personal factors. J. soc. Psychol., 1958, 48, 3-10.

13. John, E. R. Contributions to the study of the problem-solving process. Psychol. Monogr., 1957, 71, No. 447.

14. John, E. R. and Miller, J. G. The acquisition and application of information in the problem solving process. Behavioral Science, 1957, 2, 291-300.

15. Johnson, D. M. Generalization of a scale of values by the averaging of practice effects. J. exp. Psychol., 1944, 34, 425-436.

16. Johnson, D. M. The Psychology of Thought and Judgment. New York, Harper, 1955.

17. Johnson, D. M. Serial analysis of thinking. Ann. New York Acad. Sci., 1960, 91, 66-75.

18. Johnson, D. M. Formulation and reformulation of figure-concepts. Amer. J. Psychol., 1961, 74, 418-424.

19. Johnson, D. M. and Hall, E. R. Organization of relevant and irrelevant words in the solution of verbal problems. J. of Psychol., 1961, 52, 99-104.

20. Johnson, D. M., Lincoln, R. E., and Hall, E. R. Amount of material and time of preparation for solving problems. J. of Psychol., 1961, 51, 457-471.

21. Johnson, D. M. Main forms of thought. Centennial Rev., 1962, 6, 120-130.

22. Johnson, D. M. and Jennings, J. W. Serial analysis of three problem-solving processes. Unpublished paper.

23. Lewis, D. Quantitative Methods in Psychology. New York, McGraw-Hill, 1960.

24. Levy, L. H. Context effects in social perception. J. abnor. and soc. Psychol., 1960, 61, 295-297.

25. McGarvey, Hulda R. Anchoring effects in the absolute judgment of verbal materials. Arch. of Psychol., 1943, No. 281.

26. Parducci, A. and Marshall, Louise M. Context effects in judgments of length. Amer. J. Psychol., 1961, 74, 576-583.

27. Parducci, A. and Marshall, Louise M. Effects of mean, midpoint, and median upon adaptation level in judgment. J. exp. Psychol., 1961, 61, 261-262.

28. Pratt, C. C. Time-error in the method of single stimuli. J. exp. Psychol., 1933, 16, 798-814.

29. Rimoldi, H. J. A., Devane, J. and Haley, J. Characterization of processes. Ed. and Psychol. Measur., 1961, 21, 383-392.

30. Rogers, S. The anchoring of absolute judgments. Arch. of Psychol., 1941, No. 261.

31. Segall, N. H. The effect of attitude and experience on judgments of controversial statements. J. abnor. and soc. Psychol., 1959, 58, 61-68.

32. Volkmann, J. The anchoring of absolute scales. Psychol. Bull., 1936, 33, 742-743.

33. Woodworth, R. S. Experimental Psychology. New York, Holt, 1938.

34. Woodworth, R. S. and Schlosberg, H. Experimental Psychology. New York, Holt, 1954.