6-1970

Effects of proportion of positive instances and degree of restriction on the induction of a principle

Dennis C. Dervin

EFFECTS OF PROPORTION OF POSITIVE INSTANCES AND DEGREE OF RESTRICTION ON THE INDUCTION OF A PRINCIPLE

A Thesis

Presented to the

Department of Psychology

and the

Faculty of the Graduate College

University of Nebraska at Omaha

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Dennis C. Dervin

June, 1970

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Accepted for the faculty of The Graduate College of the University of Nebraska at Omaha, in partial fulfillment of the requirements for the degree Master of Arts.

Graduate Committee

[redacted]
Name Department

Gordon L. Gilbert Ed.

C. Raymond Wilhoit, Psy.

John Brown

Kenneth A. Deffenbacher
Chairman

Table of Contents

page

Introduction. 1

Method. 6

Subjects. 6

Procedures. 6

Materials 8

Hypotheses. 9

Results 10

Discussion. 14

References. 17

Appendix A. 19

Appendix B. 21

Abstract

An experiment is reported in which student subjects attempted to discover a principle by pairing numbers and letters. Seven groups of subjects varied on two dimensions, proportion of positive instances sampled and degree of restriction in choosing number-letter pairs. It was found that subjects who sampled a higher proportion of positive instances were more successful than those sampling a lower proportion. Furthermore, subjects who were unrestricted in their choice of number-letter pairs, because they sampled a higher proportion of positive instances, performed more efficiently than subjects who were restricted and sampled a lower proportion of positive instances. Finally, when both restricted and unrestricted groups were matched for proportion of positive instances sampled, performance was found to be similar. It was pointed out that there is a need in future research on problem solving to match treatment groups for proportion of positive instances sampled.

Acknowledgements

Grateful acknowledgement is extended to Dr. Kenneth Deffenbacher who served as major thesis advisor and teacher. The writer also wishes to extend acknowledgement to Dr. C. Raymond Millimet and Dr. Evan Brown, thesis committee members, for their assistance in statistical and methodological issues.

To the writer's sister, Mrs. Carol Ann Borsh, a special acknowledgement is given for her expert typing and patience. And, finally, a very special thank you must be given to the writer's wife, Rita, without whom this effort would not have been possible.

Introduction

Within the area of human verbal learning, much experimental effort has been expended on the determination of how individuals form concepts or principles in a laboratory situation. In the recent past, experimental evidence has been accumulated concerning the effects of positive instances versus negative instances on performance in concept learning (e.g., Hovland and Weiss, 1953; Haygood & Stevenson, 1967). Bourne (1966) defines a positive instance as an example of what a concept is and a negative instance as an example of what a concept is not. There has also been research devoted to the effects of the degree of restriction placed upon the S in his attempt to discover a principle (e.g., Duncan, 1964). This thesis is concerned with the relationship between these two variables, namely, positive versus negative instances and degree of restriction.

In an experiment reported by Wason (1960), Ss were required to discover a principle, which applied to instances consisting of three numbers each. The principle to be discovered was three numbers in increasing order of magnitude. Ss were shown the series 2, 4, 6, and told that this instance was an example of the principle. They were then required to discover the principle by creating more instances, utilizing information, given after each set was created, as to whether or not the set conformed to the principle. Wason reported

that the greater proportion of Ss who had difficulty discovering the principle engaged in "enumerative" rather than "eliminative" behavior. Wason defined "enumerative" behavior as the continued production of positive instances, such as 10, 12, 14 but the announcement of incorrect principles such as: series of even numbers. Ss thus perseverated in enumerating positive instances, but did not eliminate incorrect hypotheses.

In a second experiment reported by Wetherick (1962), it was found that by a slight modification of the paradigm used by Wason (1960), enumerative behavior could be reduced and the probability of S's success in discovering the principle was greatly increased. Wetherick contended that one of four situations could occur on any one trial. Either the instance could be positive or negative (with respect to the principle to be discovered); or, the instance could conform or not conform to the hypothesis currently held by S. Only two of these cases; namely, a positive instance that does not conform to S's current hypothesis, or, a ngative instance that does conform to the current hypothesis, is effective in eliminating wrong hypotheses. Wetherick required his Ss to state, after the creation of each instance, whether they thought the instance would be a positive or negative example of the principle. By doing this, E was able to establish whether

or not S was in fact eliminating a wrong hypothesis. It was possible for E to make such a judgement since, if S's prediction of the instance was positive (conforming to his current hypothesis) and E's feedback was negative (not an example of the principle), one of the two situations that Wetherick contends leads to the elimination of an incorrect hypothesis has occurred. When an instance was created that did not conform to S's current hypothesis and E's feedback was positive, the other condition which eliminates a wrong hypothesis has occurred. The results of his study indicated that enumerative behavior was in fact reduced (thus, aiding Ss in discovering the principle). However, eliminative behavior still did not occur at a very high rate. Wetherick attributed this result to the fact that whenever one of the more obvious hypotheses was under consideration by S, (e.g., 4, 8, 16: numbers which are multiples of four:) negative feedback was not delivered by E, since the instance created by S was an example of the principle. Thus, a conforming-negative instance did not arise to eliminate S's incorrect hypothesis and hence, eliminative behavior occurred at a lower rate.

Duncan (1964) noted that in the Wason study S was allowed to write down, for every instance, any three numbers that he wished. Duncan contended that this may have accounted

for the results obtained by Wason; namely, that Ss, when they are unrestricted in their choice of instances, engaged more in enumerative rather than eliminative behavior. Duncan therefore utilized four groups varying along the dimension of restriction. Essentially, Duncan's method is an attempt to account for the criticism of the Wason study raised by Wetherick. Group I was totally unrestricted in its choice of instances. Groups II and III were partially restricted; and, Group IV was totally restricted in its choice of instances. Ss' task consisted of pairing the numbers 1-8 with the letters A-J. The principle was simply that any of the four even numbers could be paired with the letters A-E, and any of the odd numbers could be paired with letters F-J. Duncan originally hypothesized that those groups which were restricted in some degree would discover the principle more quickly than the unrestricted group. It was felt that the restricted groups would have the advantage over the unrestricted group in that these Ss would be prevented from perseverating in the wrong hypothesis (as was the case in the Wason study). The results, however, did not confirm Duncan's hypotheses. The unrestricted group was found to be superior to all other groups in discovering the principle.

There may have been, however, a confounding factor in the

Duncan study. In the partially and totally restricted groups, Duncan arbitrarily imposed a restriction such that within each block of 40 trials, half the instances would be positive, and half would be negative instances of the principle. In the unrestricted group, this restriction could obviously not be utilized since Ss were free to choose any instance they wished. The proportion of positive versus negative instances, in the unrestricted group, could thus deviate significantly from the 50-50 proportion imposed upon the restricted groups. Hovland and Weiss (1953) presented evidence that positive examples of a concept are much more effective in transmitting information about that concept than are negative examples. These Es also found that mixed proportions of positive and negative examples fall somewhere between all positive series and all negative series in their effectiveness of information transmission.

More recent experimental evidence concerning the proportion of positive and negative instances, and its effect on the learning of a concept, has been reported by Haygood and Stevenson (1967). Haygood and Stevenson studied the effects of varying the proportion of positive and negative instances on the speed of learning a concept. He varied the proportion of positive from .1 to .8. The results clearly indicated that as the proportion of positive and negative instances increases, the time required

to master a concept decreases in a direct relationship.

The present study attempted to determine whether the results obtained by Duncan were in fact due only to the varying degrees of restriction placed upon his groups, or, if his results might have been at least partially due to the different proportions of positive and negative instances sampled by the unrestricted group on the one hand, and the restricted groups on the other.

Method

Subjects. Ss were drawn from an introductory psychology course at the University of Nebraska at Omaha as a part of a course requirement. Ten Ss were unsystematically assigned to each of seven treatment groups.

Procedure. In order to test the contention that the results obtained by Duncan were at least partially an artifact of the proportion of positive instances sampled (the unrestricted group sampling a higher proportion than the restricted groups), it was necessary to run the unrestricted group first. This was necessary in order to determine the average proportion of positive examples sampled by Ss in this group.

Of the seven groups, four were replicates of Duncan's conditions. That is, one group was totally unrestricted in its choice of instances. This group is referred to as the U Group. In Group RN (restricted numbers) S was allowed to choose any letter he wished, but was told which number to

write down. Similarly, Group RL (restricted letters) chose any number, but was told which letter to record. And, finally, Group RNL (restricted numbers and letters) was told which number and letter to record. All of the above three groups (RN, RL, RNL) had the 50:50 proportion of positive and negative instances imposed upon them as was the case in Duncan's experiment. These groups are referred to respectively as RN50, RL50 and RNL50.

The remaining three groups (also RN, RL, RNL) operated under the same conditions as those stated above. However, these three groups sampled the same proportion of positive instances as that sampled by the U group. Similarly, they are referred to as RN(U%), RL(U%), RNL(U%). These groups have been added in order to test the contention that Duncan's results were due in part at least to the effect of proportion of positive examples sampled.

The task was the same as that in the Duncan study. S was required to discover the principle "any of the even numbers (2, 4, 6, 8) may be paired with any of the letters A-E; and, any of the odd numbers (1, 3, 5, 7) may be paired with any of the letters F-J," by creating number-letter pairs under the conditions stated above. Before the experiment began each S was given instructions concerning the experimental task, and the method he was to employ depending upon which

group he was assigned (see Appendix for a complete set of instructions). Each trial consisted of the creation of a number-letter pair, followed by experimenter feedback as to whether the instance was or was not an example of the principle. S was allowed to work at his own pace, and was allowed to guess at the principle at any time. After each guess, E informed S whether the hypothesis was correct, partially correct or wrong. If S had not stated the principle correctly after 25 minutes, he was stopped and recorded as a nonsolver. All of the above procedures were replications of the Duncan study. In addition, S was told at the end of the experiment not to discuss the nature of the experiment with others.

Materials. The only materials required for this experiment were experimental sheets on which S recorded his number-letter choices. There were two forms of this sheet. On both forms of the sheet, the numbers 1-8 were printed horizontally in normal order in the upper left corner of the sheet. The letters A-J, were printed horizontally in normal order in the upper right corner of the sheet. One form of the sheet contained the numbers and letters printed at the top, with 80 blanks in four columns of twenty each. Groups U, RN and RL used this form. The other form, with the numbers and letters printed at the top, consisted of eighty blanks, four columns

of twenty each, with one of the numbers (1-8) printed in each blank. Within any block of eight trials, each number was used once with the only restriction that the assignment of numbers within blocks be random. Group RNL used this form.

Hypotheses. The present study formulated three major sets of hypotheses in order to test the contention that the variance obtained by Duncan was at least partially due to differences in proportion of positive instances sampled. The first set of these hypotheses predicted that, as was the case in the Duncan study, the unrestricted group (group U) would perform more efficiently than the restricted groups (which were replicates of the Duncan study: RN50, RL50 and RNL50).

The second set of hypotheses, which did not predict direction, tested whether those restricted groups (added by the present study) that sampled the same proportion of positive instances as the unrestricted group differed in performance from the unrestricted group. Since all groups in this set are matched in terms of proportion of positive instances sampled, this set of hypotheses provided an unconfounded test of the effectiveness of the restriction variable (which was Duncan's purpose).

Finally, the third set of hypotheses predicted that those restricted groups that sampled a higher proportion of positive instances would perform better than those restricted groups

sampling a lower proportion. These hypotheses were in keeping with experimental evidence cited previously (e.g., Hovland and Weiss, 1953; Haygood and Stevenson, 1967). This set of hypotheses provided a test of whether the performance of Duncan's restricted groups was in fact impaired by the arbitrarily low proportion of positive instances which they sampled.

Results

As was mentioned previously, in order to match the three restricted groups, added in the present study, with the unrestricted group in terms of positive instances sampled, it was necessary to run the U group first. Upon examination of the data from this group, it was found that the average proportion of positive instances sampled by Ss in this group was 0.70. Accordingly, the experimental conditions of the three restricted groups added in this study, RN(U%), RL(U%), and RNL(U%), were such that within a block of 40 trials, 70% of the instances were positive. Since, however, many Ss in these three groups reached solution prior to trial 40, the possibility existed that they were actually sampling proportion of positive instances different than the intended 0.70. In order to test this possibility, a z-test for the difference between observed and predicted proportions was applied to the average proportion of positive instances sampled within each of the three groups (Spiegel, 1961). For the three groups, RN(70%), RL(70%) and RNL(70%), the following two-tailed probability levels

associated with the observed z-values were found respectively: .4502, .1052 and .4502. Since none of these probability levels even remotely approached conventional significance levels, it was concluded that the three groups were in fact sampling the intended proportions of positive instances.

Those Ss who failed to reach solution within the required time limit were not included in the final data analysis. One S was classified as a non-solver in each of the following groups: U, RN(70%), RL(70%), and RNL(70%). This constituted a proportion of 0.90 of Ss reaching solution in those groups sampling a higher proportion of positive instances. Four Ss were classfied as non-solvers in group RN50, and groups RL50 and RNL50 had three non-solvers each. The proportion of Ss reaching solution in these groups were 0.60, 0.70, and 0.70 respectively. In the Duncan study (which included only groups U, RN50, RL50, and RNL50), the proportion of solvers were: 0.86, 0.52, 0.54, and 0.74. In the present study it was decided to determine whether there was an extraneous variable operating that could account for the differential drop-out rate among the groups (0.90 of those groups sampling a higher proportion of positive instances solved the problem, while only 0.66 of those Ss sampling the lower proportion reached solution). In order to achieve this, a large sample binomial test was applied to the proportion of Ss within each group who solved the task, (Siegel, 1956). This procedure

yielded a z-value of 1.10 ($p = .1357$), which was non-significant. It was concluded that no extraneous variable was operating to influence the differential drop-out rate among the groups.

The data obtained in this study did not permit analysis by parametric techniques, as there was truncation of the data imposed by the 25 minute solution time requirement. The median number of pairs to solution for the Duncan and the present study are reported below.

Groups	Duncan	Dervin
U	39.5	34.0
RN50	92.0	54.5
RL50	82.0	47.0
RNL50	60.0	51.0
RN70	----	20.0
RL70	----	33.0
RNL70	----	34.0

An overall analysis of the data was carried out using the Kruskal-Wallis Analysis of Variance by Ranks on number-letter pairs to solution, ($H=55.70$ with 6 d.f., $p<.001$). In an attempt to isolate differences among the groups, based on the above result, three separate Kruskal-Wallis tests were performed. Each of these three tests was applied to the groups in the three sets of planned comparisons. An analysis involving groups U, RN50, RL50, and RNL50 yielded an H of 16.02 (d.f. = 3, $p<.01$). Again, it can be asserted that these four groups were different in performance levels. Analysis involving groups

U, RN(70%), and RNL(70%) produced an H of 6.58 (d.f. = 3, $p<.10$). Apparently, then, these groups were not significantly different in terms of performance. Finally, the analysis of groups RN(70%), RL(70%), RNL(70%), RN50, RL50, and RNL50 yielded an H of 22.0 (d.f. = 5, $p<.001$).

On the basis of the results, it was decided to proceed with the three sets of planned comparisons, using the Mann-Whitney U test for independent samples. This test is analogous to the t-test for uncorrelated data.

All three comparisons in the first set were significant. The difference between U and RN50 yielded a U value of 9.0 ($n1 = 6$, $n2 = 9$, $p<.025$). The U versus RL50 comparison provided a U value of 12.0 ($n1 = 7$, $n2 = 9$, $p = .025$). Finally, the difference between U and RNL50 yielded a U value of 8.0 ($n1 = 7$, $n2 = 9$, $p<.01$). Given the results of the above three comparisons, it was concluded that the U group (as Duncan found) performed significantly better than those restricted groups which sampled a proportion of positive instances equal to .50.

The second set of comparisons yielded only one comparison which even approached significance. The comparison between U and RN(70%) produced a U value of 19.0 ($n1 = 9$, $n2 = 9$, $p<.10$). On the basis of this set of comparisons, it was concluded that the U group and groups RN70, RL70, and RNL70 were essentially similar in terms of performance.

All three comparisons in the third set were significant. The comparison between RN(70%) and RN50 yielded a U of 2.0 ($n1 = 6$, $n2 = 9$, $p = .001$). The RL(70%) versus RL50 comparison produced a U of 9.5 ($n1 = 7$, $n2 = 9$, $p < .025$). Finally, the difference between RNL(70%) and RNL50 yielded a U of 8.5 ($n1 = 7$, $n2 = 9$, $p < .01$). These results corroborated the prediction that those groups sampling a higher proportion of positive instances would perform better than those sampling lower proportions.

Discussion

The present study replicated the results obtained by Duncan. That is, the U group was found to be superior to those restricted groups sampling a 0.50 proportion of positive instances. Furthermore, those restricted groups which sampled a 0.70 proportion of positive instances were shown to be superior to the restricted groups sampling a 0.50 proportion. The attempt to demonstrate differences between the unrestricted and restricted groups matched for proportion of positive instances sampled failed to produce significant results.

On the basis of the results of the present study, it appears reasonable to conclude that the effects obtained by Duncan were confounded by the operation of the varying proportion of positive instances sampled by his groups. The results of the separate Kruskal-Wallis analysis and the second set of

comparisons suggests that the restriction variable did not account for a significant proportion of the variance. However, in a study by Hunt (1965), the restriction variable was found to be a significant one. In the Hunt study, Ss were required to learn an artificial language, involving letter strings, under either of two conditions, a selection or a reception paradigm. Essentially, these two conditions are similar to restricted and unrestricted conditions; selection being unrestricted and reception being restricted. From his data, Hunt concluded that those Ss who were free to construct and test their own instances were more efficient in solving the task. Since the results of the present study and those obtained by Hunt are directly opposed, it appears that the effect of the restriction variable is not generalizable to all experimental tasks. The task performed by Ss in the Hunt experiment appears to be considerably more difficult than that performed by Ss in the present study. It would seem then that the restriction variable is an effective variable in difficult tasks, but fails to have a significant effect in easier tasks as the one utilized in the present study. Two additional studies, Huttenlocher (1962) and Murray and Gregg (1969), in which Ss learned a single cue concept (again the experimental task utilized in these studies appears to be easier than that of Hunt's) not only failed to demonstrate the superiority of the performance of the unre-

stricted Ss, but found that restricted Ss reached solution significantly faster. The restriction variable is apparently task bound.

A further need also became apparent in the course of the present study. It would have been desireable to be able to state confidence intervals for differences between median pairs to solution among the various groups in order to establish intervals of true differences among these groups. At the present time, however, no statistical procedure exists that allows such a statement. Since much of the data obtained by psychologists is non-parametric in nature, the need for the development of such a procedure is overdue.

Finally, since those restricted groups sampling a 0.70 proportion of positive instances were shown to be superior to the restricted groups which were replicates of the Duncan study, it appears that the major portion of the variance found in the Duncan study was a result of the positive instance variable and not the restriction variable as Duncan concluded.

Future research in the area of problem solving using a restriction variable should make every effort to match treatment groups for proportion of positive instances sampled.

References

Bourne, L.E., Jr. Human conceptual behavior. Boston: Allyn and Bacon, 1966.

Duncan, D.P. Induction of a principle. The Quarterly Journal of Experimental Psychology, 1964, 16, 373-377.

Haygood, R.C., & Stevenson, M. Effects of proportion of positive instances upon concept learning. Psychological Reports, 1967, 20, 179-82.

Hovland, C.I., & Wriss W. Transmission of information concerning concepts through positive and negative instances. Journal of Experimental Psychology, 1953, 45, 175-82.

Hunt, E.B. Selection and reception conditions in grammar and concept learning. Journal of verbal learning and verbal behavior, 1965, 4(3), 211-215.

Huttenlocker, J. Effects of manipulation of attributes on efficiency of concept formation. Psychological Reports, 1962, 10, 503-509.

Seigel, S. Nonparametric statistics for the behavior sciences. New York: McGraw-Hill, 1956.

Spiegel, M.R. Theory and problems of statistics. New York: McGraw-Hill, 1961.

Wason, P.C. On the failure to eliminate hypotheses in a conceptual task. The Quarterly Journal of Experimental Psychology. 1960, 12, 129-40.

Wetherick, N.E. Eliminative and enumerative behavior in a conceptual task. The Quarterly Journal of Experimental Psychology, 1962, 14, 246-9.

Appendix A
Instructions

See these numbers and letters? They have been paired according to a certain principle. I want you to try to discover what that principle is. This is the way we will do it.

RL

In each of the blanks, I want you to write down any number you see at the top of the page. After you have done this, state the number aloud. I will then tell you which letter to write down in that blank. I will then tell you whether the number-letter pair is or is not an example of the principle I have in mind.

After each pair that is an example of the principle, make a check mark, so that you will remember what an example of the principle looks like.

RN

In each of the blanks, I want you to write down any letter that you see at the top of the page. After you have done this, state the letter aloud. I will then tell you which number to write down in that blank. I will then tell you whether the number-letter pair is or is not an example of the principle I have in mind.

After each pair that is an example of the principle, make a check mark, so that you will remember what an example of the principle looks like.

RNL

In each of the blanks, you will see that there is a number written down. I want you to state that number aloud. I will then tell you which letter to write down after the number. I will then tell you whether the number-letter pair is or is not an example of the principle I have in mind.

After each pair that is an example, make a check mark so you will remember what an example of the principle looks like.

As soon as you have any idea as to what the principle is,

tell me. If you are wrong or only partly right, I will tell you so, and we will continue with the pairs until you want to make another guess as to what the principle is. Try to discover the principle as soon as possible. As soon as you discover the principle, the experiment is finished.

Appendix B
Data Sheets

2 3 4 5 6 7 8 A B C D E F G H I J

21 41 61

22 42 62

23 43 63

24 44 64

25 45 65

26 46 66

27 47 67

28 48 68

29 49 69

30 50 70

31 51 71

32 52 72

33 53 73

34 54 74

35 55 75

36 56 76

37 57 77

38 58 78

39 59 79

40 60 80

2 3 4 5 6 7 8 A B C D E F G H I J

7	21 3	41 5	61 1
4	22 4	42 7	62 4
2	23 7	43 4	63 5
1	24 5	44 6	64 7
3	25 7	45 3	65 3
6	26 8	46 1	66 1
5	27 4	47 8	67 4
8	28 3	48 2	68 6
7	29 1	49 2	69 5
5	30 5	50 3	70 8
6	31 6	51 5	71 7
2	32 2	52 7	72 2
4	33 1	53 6	73 1
3	34 3	54 8	74 3
1	35 6	55 1	75 7
8	36 8	56 4	76 8
8	37 7	57 8	77 6
6	38 2	58 2	78 5
2	39 4	59 6	79 4
1	40 5	60 3	80 2