

## Learning conceptual rules: III. Processes contributing to rule difficulty\*

HARRIET SALATAS and L. E. BOURNE, JR.  
*University of Colorado, Boulder, Colorado 80302*

Eight unique bidimensional conceptual rule forms were compared in a rule learning paradigm. It was predicted that rule difficulty order and error distributions across stimulus class defined by a truth table would be a function of Ss' preexperimental experience predominantly with conjunctive concepts. The hypothesis stated specifically that loci of difficulty with rules derive from: (a) assignment of TT stimuli (those with both relevant attributes) to the negative category; (b) assignment of FF stimuli (those with neither relevant attributes) to the positive category; (c) assignment of TT and FF stimuli to the same category. The main effects of rule and truth table class and the interaction of rule and truth table class were statistically significant. The results confirm all predictions about rule order and error distributions.

A nominal concept can be characterized as a rule-form relationship among a set of perceptible stimulus attributes. When the number of attributes is limited to two, a calculus of 16 different rule forms and corresponding class concepts is generated. Of this set, only eight are nontrivial and bidimensional. These eight rule forms fall into complementary pairs; instances which are positive for one are negative instances of its complement, and vice versa (Haygood & Bourne, 1965; Bourne, 1970). Thus, in a sense, the complete calculus can be reduced to and represented by four primary bidimensional concepts, distinguished by the rule which combines relevant attributes. These primaries are called the conjunctive, the disjunctive, the conditional, and the biconditional rules and are illustrated, along with their complements, in Table 1.

Initial empirical work has established a difference in difficulty among these rules. When the problem is to learn which one of the several rules defines a particular concept, attributes of the concept having been given during preliminary instructions, the typical order of rule difficulty (Bourne, 1970) is conjunctive (easiest), disjunctive, conditional, and biconditional (most difficult). The order might be different in attribute identification tasks, where the rule is known and the attributes must be discovered, for, under some conditions, it has been observed that biconditionals are no more difficult than disjunction (Taplin, 1971). In rule learning tasks, however, the order given above has been replicated on at least a dozen different occasions (e.g., Neisser & Weene, 1962; Bourne, 1970; Bourne & Guy, 1968a, b; Dodd, 1967).

Bourne (1970) examined several interpretations of

rule difficulty, including hypotheses based on (a) simple S-R logic, (b) the assumption of a positive focusing strategy (Bruner, Goodnow, & Austin, 1956), (c) the assumption of a strategy focusing on the smaller, more homogeneous category of stimuli, positive or negative (Bourne & Guy, 1968a, b), and (d) differences among concepts in structural complexity (Neisser & Weene, 1962). None of these theories generates the order of rule difficulty observed in empirical studies.

More recently, Sawyer and Johnson<sup>1</sup> have developed an account of rule difficulty which is consistent with available findings. In a rule learning problem, S must learn to categorize stimuli from four different classes, defined in terms of the presence (T) and absence (F) of the two relevant attributes. As in a bidimensional truth table, these stimulus classes are described as TT, TF, FT, and FF. In their theory, Sawyer and Johnson assume that the performance of naive Ss is governed by predominant extraexperimental experience with conjunctive concepts. They argue, therefore, that the primary source of errors in rule learning problems is confusion among stimuli having common F values which are assigned to different response categories. The difficulty of this process is assumed to be magnified when TT and FF stimuli—stimuli with no common relevant attributes—are both assigned to the category of positive instances. Following this logic, conjunctive concepts are clearly the easiest, for TT and FF stimuli are assigned to different categories and all stimuli sharing F values, i.e., TF, FT, and FF, are assigned to the same category (negative). Disjunctions are more difficult because TF and FT stimuli, sharing common F values with FF stimuli, are assigned to a different response category, TF and FT positive and FF negative. Conditionals and biconditionals also require that the stimuli with common F values be placed in different categories, and the effects are compounded by the necessity to place TT and FF stimuli together in the positive category. Both TF and FT stimuli are assigned differently than FF stimuli in the biconditional, whereas

\*This research was undertaken within the Institute for the Study of Intellectual Behavior and is Publication Number 43 of the Institute. The work was supported by a research grant, MH 14314, and by a Research Scientist Award, 1-K5-MH37497, both from the National Institute of Mental Health, and a research grant, GB-340-77X, from the National Science Foundation. The authors wish to acknowledge the critical but useful comments of John E. Taplin and Peder J. Johnson on an earlier version of this paper.

Table 1  
Four Primary Bidimensional Rules and Their  
Complements Represented in Truth Table Form

Primary Rules	Conjunctive	Disjunctive	Conditional	Biconditional
TT	+	+	+	+
TF	-	+	*	*
FT	-	+	+	*
FF	-	-	+	+
Difficulty Value	0	4	12	16
Complementary Rules	Alternative Denial	Joint Denial	Exclusive	Exclusive Disjunctive
TT	*	*	*	*
TF	+	*	+	+
FT	+	*	-	+
FF	+	+	-	-
Difficulty Value	5	9	6	10

only TF stimuli are differentially assigned in the conditional. Therefore, biconditionals are more difficult than conditional concepts. Stimulus classes whose assignments represent a source of difficulty within each primary rule are marked with asterisks in Table 1.

All the evidence offered by Sawyer and Johnson in support of their position was collected in experiments using the four primary bidimensional rules, whose difficulty order the theory was tailor-made to reflect. A stronger test of their position would be an application of the theory to other rules within the same logical system. In other words, can the arguments of Sawyer and Johnson be used to forecast the difficulty of complementary rules, which are mirror images of the primary bidimensional rules?

In order to extend the theory to other rules within the same logical system, certain assumptions have to be made. First, Sawyer and Johnson assert that TT instances offer no particular source of difficulty in rule learning problems because they exhibit both relevant attributes, which have been given to S by instruction, and they are placed, consistent with S's expectations, in the positive category. This assertion is true of primaries but not of complements. For all complements, TT instances are negative. Being inconsistent with S's expectations, the placement of TTs is an added source of difficulty in complementary rules and leads us to expect, all other things equal, that complements will be more difficult than primaries. Second, Sawyer and Johnson assert that the necessity to combine TT and FF instances into the positive category of a concept adds to the difficulty of the problem. Such an assignment never occurs among the complementary rules, which may lead one to expect the complements of conditional and biconditional concepts (namely, exclusion and exclusive disjunction, respectively) to be relatively easy. These complements do, however, require Ss to combine TT and FF instances into the *same* response category (in this case, negative), which might be just as difficult as combining them in the positive category.

From these considerations, we arrive at a set of assumptions which will allow us to predict the relative order of rule difficulty for all rules within the logical system. (A) S will place TT instances, instances with both given relevant attributes, in the positive category. (B) S will place FF instances, instances with neither relevant attribute, in the negative category. (C) S will place TF and FT instances in the category assigned to FF instances, whether that category is positive or negative. (D) S will place TT and FF instances in different response categories.

These assumptions describe a set of initial states or processes on information that specify response output to any stimulus. In principle, these processes should be quantifiable. In fact, however, because we have no rational model at this time, it is unclear which of several possible quantification schemes one might use. As a simple first approximation, we propose the following. A violation of any of the foregoing assumptions results in some increment of problem difficulty. The increment of difficulty is tentatively assumed to be equivalent to the number of unique stimulus patterns which must be assigned to a response category inconsistent with a given process. To see how this reasoning works, consider first a conjunctive problem. Because TTs are positive, FFs are negative, TFs and FTs are assigned in a way consistent with FFs, and there is no necessity to combine TTs with FFs, the difficulty value of conjunction is zero. The conjunctive concept is entirely consistent with S's preexperimental expectations, and he should make none or at most a few errors on that type of problem. Consider the disjunction. TTs are positive and FFs are negative. However, TFs and FTs are both assigned to the positive category, consistent with the assignment of FFs. In a stimulus population defined by two values on each of the two relevant dimensions, there is only one stimulus corresponding to each truth table class. The difficulty of the disjunctive problem, therefore, would be 2 (TF = 1 and FT = 1) relative to 0 for conjunctions. In fact, in the problems used in most rule learning experiments, there are typically three values per relevant dimension, thereby giving the disjunction a difficulty value of 4 (TF = 2 and FF = 2). Next, consider the conditional, where TTs, FTs, and FFs are positive and TFs are negative. In this case, 4 + 2, or 6, unique stimulus patterns are in violation of S's assumed informational processes. Furthermore, TT and FF instances must be placed into the same (positive) category, which magnifies the difficulty of the problem by, say arbitrarily, a factor of 2. Thus, the difficulty value for conditional problems is 12.

Similar reasoning can be applied to all rules, primary and complementary, outlined in Table 1. The analysis yields a difficulty value for each of the rules, which can then be checked against empirical data. We present these values in Table 1. Clearly, we do not anticipate the numerical values obtained by an arbitrary calculational process to correspond exactly with the quantitative data one obtains from a rule learning experiment. Still, if this

kind of reasoning is in any sense correct, the order of rule difficulty predicted ought to correspond with the empirical order. On this reasoning, then, we expect the following order of rule difficulty: conjunctive (0), disjunction (4), alternative denial (5), exclusion (6), joint denial (9), exclusive disjunction (10), conditional (12), and biconditional (16). It is the purpose of this experiment to collect data on the eight primary and complementary bidimensional rules for purposes of comparison with this predicted order.

There is a variety of alternative possible outcomes of this comparison among rules. Perhaps the simplest is based on the fact that complements are the mirror images of primary rules. All stimuli which are assigned to the negative category of a primary rule are positive instances of its complement. In a sense, the labels, positive and negative, are arbitrary. If S treats them in that way, then the order of complements ought to be exactly the same as the order of primaries; a primary ought to be essentially equivalent to its complement in difficulty. There is some evidence to support this conjecture. In a single comparison of complementary pairs, Haygood and Bourne (1965) were unable to demonstrate a statistically significant difference in difficulty. These data do not represent a stringent test, however, because the number of Ss was small and there was a numerical difference between primary and complement. Furthermore, the evidence reported by Laughlin (1968) and Giambra (1971), although collected in attribute identification problems, suggests clear and reliable differences between certain rules and their complements.

## METHOD

### Subjects and Design

The Ss were 96 college students, 36 volunteers and 60 paid recruits, none of whom had previously participated in a concept learning experiment. They were assigned in order of appearance to one of eight conditions. Four Ss did not reach the problem solving criterion in 160 trials and were replaced.

The experimental design was an 8 by 2 factorial incorporating rule learning problems based on eight different rules (four primaries and their complements) and two different pairs of relevant attributes. The primary rules were conjunctive, disjunctive, conditional, biconditional; their respective complements were alternative denial, joint denial, exclusion, and exclusive disjunction. The two pairs of relevant attributes were three-yellow and large-triangle.

### Materials and Apparatus

The stimuli were slides of four-dimensional geometric patterns with three values on each dimension. The dimensions were color (blue, yellow, or red) and size (large, medium, or small), form (hexagon, square, or triangle) and number of figures (one, two, or three). The constraints placed on the 40 slide sequence for each problem were that it contain 10 slides from each truth table class (TT, TF, FT, FF), that the first four slides include one example of each of the four truth table classes, and that the first slide be a TT instance.

The slides were projected one at a time onto a screen before S with a Kodak Ektagraphic slide projector. There was a response panel in front of the screen with two buttons, designated "+"

Table 2  
Mean Trials and Errors for the Eight Bidimensional Rules

Primary Rules	Conjunction	Disjunction	Conditional	Biconditional
Mean Trials	6.92	8.83	60.58	66.42
Mean Errors	2.67	3.33	23.08	27.25
Obtained Order	1	2	7	8
Predicted Order	1	2	7	8
Complementary Rules	Alternative Denial	Joint Denial	Exclusion	Exclusive Disjunction
Mean Trials	35.00	49.17	32.50	40.33
Mean Errors	13.75	21.42	12.67	18.17
Obtained Order	4	6	3	5
Predicted Order	3	5	4	6

and "-" Informative feedback was given by a light on the correct response button. Responses made by S and the appropriate feedback were recorded on an Esterline-Angus operations recorder. LVE solid state programming modules were arranged to provide for an immediate feedback signal of 1-sec duration after each of S's responses and for a 5-sec intertrial interval before the next slide appeared. The slide tray was repeated after the 40-trial run if necessary.

### Task and Procedure

The Ss were required to classify a series of stimulus patterns into two categories, positive and negative instances of a concept. Correct classification for any stimulus was determined both by the pair of attributes relevant in the particular problem and by the conceptual rule which specified the relationship between these attributes. In each problem two stimulus dimensions were irrelevant to solution.

The instructions to S included a brief description of a simple nominal concept. Three cards which exemplified all the possible values on the four stimulus dimensions were placed before S and described. In order to familiarize S with the ideas of rule and relevant attributes of a concept, he was required to sort playing cards into four piles on the basis of two attributes, red-black and face-nonface cards. The required piles were defined by the four truth table classes. After this exercise, S was told the names of the two relevant attributes for his experimental problem and was instructed how to respond to the stimuli, with particular emphasis on the positive-negative distinction between responses. All problems were self-paced, S being allowed as much time as needed to make any response. The experiment was terminated when S had reached a criterion of 12 consecutive correct responses or reached 160 trials without solving. Of the four Ss who did not solve, three were assigned to a biconditional problem and one to a conditional problem.

## RESULTS

Separate analyses of variance were performed on errors and trials to criterion. The analyses were essentially the same statistically, whether nonsolvers (with arbitrary scores of 160 trials and 80 errors) or their replacements were included. Therefore, only the results from solvers are presented.

Table 2 presents the mean trials and the mean errors to solution for each of the eight rules. Analysis of variance revealed differences among rules for the trials data [ $F(7,80) = 4.84, p < .001$ ]. Neither problem nor the Problem by Rule interaction was significant in either analysis. F tests based on individual dfs showed significant differences ( $p < .01$ ) between the following

Table 3  
Error Distributions Across Truth Table  
Categories for the Eight Rules

Primary Rules	Conjunction	Disjunction	Conditional†	Biconditional†
TT	.33	.42	1.25	1.75
TF	1.08	1.25*	7.50*	8.08*
FT	.92	.67*	2.42	6.08*
FF	.33	1.00	8.92*	11.33*
Complementary Rules	Alternative Denial	Joint Denial	Exclusion†	Exclusive Disjunction†
TT	5.58*	4.92*	5.17*	5.33*
TF	2.83	6.50*	3.25*	4.25*
FT	2.67	3.50*	2.58	5.17*
FF	2.67*	6.50*	1.67	3.42

\*Stimulus classes which the theory predicts will be relatively difficult for S to handle when the indicated rule applies.  
†Rules for which classes to be predicted to be more difficult produced significantly ( $p < .05$ ) more errors than classes predicted to be easier.

sets of rules: (conjunctive, disjunctive), (exclusion, alternative denial, exclusive disjunction, joint denial), and (conditional, biconditional). Within the second cluster, exclusion differed from joint denial and exclusive disjunction ( $p < .05$ ) and alternative denial differed from joint denial ( $p < .05$ ). More importantly, the order of rule difficulty is predicted accurately by the modified Sawyer-Johnson model, except for exclusion/alternative denial and joint denial/exclusive disjunction reversals within the second rule cluster. The alignment is consistent with predictions by a statistical test of ordered hypotheses,  $L = 2215.5$ ,  $p < .01$  (Page, 1963).

Table 3 shows the distribution of errors on each truth table stimulus class within each rule. Rule [F(7,80) = 5.10,  $p < .001$ ], truth table class [F(3,240) = 6.88,  $p < .001$ ], and their interaction [F(21,240) = 6.41,  $p < .001$ ] were all statistically reliable. Truth table classes which are predicted, on the basis of the Sawyer-Johnson theory, to be especially difficult within each rule are distinguished by asterisks. The difficulty patterns revealed correspond reasonably well with expectations from the theory. F test comparisons of difficult vs easy classes were made within all rules except conjunction and joint denial. All differences were in the predicted direction and were statistically significant ( $p < .05$ ) except in disjunctive and alternative denial problems. Performance was so near perfect on disjunctions that interclass differences were negligible. In alternative denial problems, the numerical difference was in accord with predictions, but an unexpected small number of errors on FF instances eliminated any possibility of statistical significance.

The revised Sawyer-Johnson model used to predict the order of rule difficulty in this experiment rests on a number of testable assumptions. The tests performed are summarized in Table 4. They confirm, without exception, the following assertions of the theory. (A) Ss

place instances with both given relevant attributes in the positive category. A comparison of the number of errors made on TT instances when they are assigned to the positive in contrast to the negative category was statistically significant ( $F = 34.20$ ,  $p < .01$ ). (B) Instances of the FF class will be placed in the negative category. A comparison of errors made on FF instances when the rule assigns them to the positive as opposed to the negative category was statistically significant ( $F = 60.80$ ,  $p < .01$ ). (C) Instances with one but not the other given relevant attribute will be categorized with FF instances, whether positive or negative. Because TF and FT instances are not always assigned in the same way by the rules of the system, separate statistical tests were performed. In both cases fewer errors were made on FT or TF instances (TF:  $F = 4.64$ ,  $p < .025$  and FT:  $F = 5.37$ ,  $p < .025$ ). (D) Finally, the model assumes that all rules which assign TT and FF instances to the same category, either positive or negative, will be more difficult than rules in which these classes are assigned differently. This assumption was evaluated in two ways. The two sets of rules were compared in total errors to solution ( $F = 157.36$ ,  $p < .01$ ) and in terms of the number of errors made on TT and FF instances only ( $F = 33.55$ ,  $p < .01$ ). Arbitrarily, the difficulty factor introduced by combining TT and FF instances into the same category was assumed to have a doubling effect. While it may be fortuitous, the means upon which both of the preceding tests are based differ approximately by a factor of two.

## DISCUSSION

Sawyer and Johnson<sup>1</sup> proposed an explanation of conceptual rule difficulty based primarily on assumptions about S's preexisting extralaboratory habits. In the present study, it has been shown that an elaboration of the Sawyer-Johnson model predicts not only the order of difficulty among the primary rules but is also consistent with the ordering of complementary rules. The success of the extension is not trivial, for the difficulty order which applies to complements is not identical to that of their primaries.

The amended theory makes the following assumptions which are verified by the details of this experiment. (1) Any stimulus that contains the two given relevant

Table 4  
Mean Errors in Particular Truth Table Stimulus Classes When Assumptions of the Model Are or Are Not Violated

Assumption	Not		F	p <
	Violated	Violated		
TTs are +	21.00	3.75	34.20	.001
FFs are -	29.42	6.42	60.80	.001
TFs With FFs	5.14	1.96	4.64	.025
FTs With FFs	15.42	8.58	5.37	.025
TTs and FFs Separate:				
TT and FF Classes Only	38.83	21.75	33.55	.001
All Classes	78.17	41.17	157.36	.001

attributes of the concept (TT instances) is assigned to the positive category. Ss commit roughly five times as many errors on TT instances when they are negative as when they are positive. (2) Any stimulus that lacks both of the given relevant attributes is placed in the negative category. In this experiment, five times as many errors were committed on FF-positive assignments as on FF-negative assignments. (3) Stimuli with one but not both of the given relevant attributes (TF and FT instances) are placed together into the category assigned to FF instances. About twice as many errors were made on TF and FT instances when their assignments differed from FF instances than when their assignments were identical to FF instances.

This experiment, therefore, provides considerable empirical support for the assumptions of the theory. The conceptual basis of these assumptions derives primarily from what Bruner, Goodnow, and Austin (1956) have called the conjunctive set, which controls the performance of naive Ss in conceptual problem solving tasks. The general idea is that, given no information to the contrary, Ss assume the unknown concept to be a conjunction. Given the two relevant attributes of the concept, S immediately generates a set of stimulus-response assignments consistent with a conjunctive rule. Information provided in the context of the experiment which controverts his set requires S engage in additional information processing activities, which logically will take several trials to complete. The difficulty of processes required by nonconjunctive rules is not certain, either on logical or intuitive grounds. In an effort to develop a metric, we have arbitrarily assumed that the difficulty is at least partly related to the number of stimuli that S must learn to assign to a conflicting response category. In the stimulus population used in this experiment, each of the relevant dimensions has three values. There are, therefore, nine unique combinations of values from the two relevant dimensions. Across the classes of the truth table, these nine combinations distribute themselves in a 1 : 2 : 2 : 4 ratio for TT, TF, FT, and FF stimuli, respectively. Thus, when a TF stimulus must be assigned to a conflicting response category, we assume that two units of difficulty are engendered. When the FF class is assigned to a conflicting category, four units of difficulty are generated. While this assumption is arbitrary, in fact, it fits the empirical results.

A further test of the assumption that difficulty is in part controlled by the number of stimuli whose assignment conflicts with a conjunctive set derives from changing the intradimensional variability of the stimulus population. Suppose, for example, stimulus dimensions have two rather than three values. Then the distribution

of stimuli across truth table classes is 1 : 1 : 1 : 1. Suppose the stimuli have four values per dimension; the distribution is 1 : 3 : 3 : 9. It is easy to show that these distributions lead to different orders of rule difficulty. Therefore, both the generality of the revised Sawyer-Johnson model and the assumed importance of number of stimuli per truth table class within that theory can be tested by varying intradimensional variability.

Finally, we should note that the proposed model applies only to the performance of naive Ss, that is, to the performance of Ss who have had no preliminary rule learning training in the laboratory. It is only with these Ss that we can reasonably assume conjunctive set toward the concept problems. It is well known that massive intra- and interruler transfer effects exist in the rule learning paradigm (Bourne, 1970; Bourne & Guy, 1968a). These transfer effects lead to definite reorderings of rule difficulty. Indeed, with a well practiced S, there are *no* differences in rule difficulty (Bourne, 1970), and a different kind of model, based on a truth table approach to these problems, is indicated.

## REFERENCES

- Bourne, L. E., Jr. Knowing and using concepts. *Psychological Review*, 1970, 77, 546-556.
- Bourne, L. E., Jr., & Guy, D. E. Learning conceptual rules: I. Some interruler transfer effects. *Journal of Experimental Psychology*, 1968a, 76, 423-429.
- Bourne, L. E., Jr., & Guy, D. E. Learning conceptual rules: II. The role of positive and negative instances. *Journal of Experimental Psychology*, 1968b, 77, 488-494.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. *A study of thinking*. New York: Wiley, 1956.
- Dodd, D. H. Rule learning transfer effects. Unpublished doctoral dissertation, University of Colorado, 1967.
- Giambra, L. M. Selection strategies for eight concept rules with non-exemplar start cards. *Journal of Experimental Psychology*, 1971, 87, 78-92.
- Haygood, R. C., & Bourne, L. E., Jr. Attribute and rule learning aspects of conceptual behavior. *Psychological Review*, 1965, 72, 175-195.
- Laughlin, P. N. Focusing strategy for eight concept rules. *Journal of Experimental Psychology*, 1968, 77, 661-669.
- Neisser, U., & Weene, P. Hierarchies in concept attainment. *Journal of Experimental Psychology*, 1962, 64, 640-645.
- Page, E. B. Ordered hypotheses for multiple treatments: A significance test for linear ranks. *American Statistical Association Journal*, 1963, 24, 216-230.
- Taplin, J. E. Effect of initial instance on attribute identification of concepts using a selection procedure. *Journal of Experimental Psychology*, 1971, 88, 177-181.

## NOTE

1. The Sawyer-Johnson model is as yet unpublished. Details can be found in a paper presented at the Rocky Mountain Psychological Association convention in Santa Fe, in 1970, by C. R. Sawyer and P. J. Johnson of the University of New Mexico, entitled "A concept learning model" and in a PhD dissertation presented by C. R. Sawyer in 1972 to the University of New Mexico, entitled "A conceptual rule learning model."

(Received for publication November 20, 1973;  
accepted January 3, 1974.)