

Transitions Between Modes of Inquiry in a Rule Discovery Task

Naftali Halberstadt¹ and Yaakov Kareev

Goldie Rotman Center for Cognitive Science and Education
School of Education
The Hebrew University of Jerusalem
Jerusalem 91905, Israel.
Email: BAUHN@VM1.HUJI.AC.IL

Abstract

Studies of rule discovery behavior employ one of two research paradigms: In the reception paradigm the item evaluated on each trial is provided by the researcher; in the selection/generation paradigm the item to be evaluated is selected or generated by the subject. The prevalence of both paradigms and their correspondence to well established modes of scientific inquiry led us to the hypothesis that if given the choice, subjects would employ both modes of inquiry. To test this hypothesis 27 adults and 27 8th graders solved three rule discovery problems in a computer environment which allowed free transitions between item reception and generation. Almost all the adults and roughly half the children employed both modes of inquiry on at least one problem, with adults much likelier to generate items. The use of a method of inquiry came in blocks with generation tending to follow reception. An inverse relationship was found between item generation and the proportion of positive instances supplied by the environment. Within both age groups, consistent individual differences were found regarding inquiry style. These results shed new light on inquiry behavior and demonstrate the desirability of letting subjects freely choose between differing modes of inquiry.

Theories which attempt to explain behavior in rule discovery tasks

usually adopt one of two orientations regarding the process: Data driven or theory driven. This differentiation has also been characterized as "bottom-up" or "top-down."

Most empirical studies of the inquiry process, have implicitly adopted one of these two approaches, in that they employed one of two research paradigms: The reception paradigm, or the selection paradigm.

Under the reception paradigm the stimuli to be classified are presented by the experimenter in a pre-arranged order, giving the experimenter complete control over the stimuli encountered by the subject. Under the selection paradigm the subject is presented with the entire array of stimuli at the outset and then freely chooses stimuli in order to discover the concept. Under the closely related generation paradigm the subject creates or generates the instances to be tested in the search for the concept or rule. The advantage of this mode of inquiry is that it is relatively easy to infer the subject's thought processes given the item chosen and the current hypothesis.

The modern study of rule-discovery and hypothesis testing started with Bruner, Goodnow and Austin (1956) and continued to generate much interest and activity throughout the years (for recent studies and theoretical developments see Evans (1989), Klahr and Dunbar (1988), Klayman and Ha (1987, 1989). Two findings regarding rule discovery behavior are particularly relevant to us. First, solving such tasks subjects exhibit a

¹ This paper is based on data collected for a doctoral dissertation submitted to the Hebrew University of Jerusalem by the first author. The work was supported in part by grants from the Israel Foundations Trustees and The NCJW Research Institute for Innovation in Education.

strong confirmation bias (Wason, 1960) which as been found very difficult to modify (Mynatt, Doherty and Tweney, 1978; Wason and Johnson-Laird, 1972). More recently Evans (1989) offered an alternative perspectives on the phenomenon while Klayman and Ha (1987) view the search for positive instances of a target category as part of an efficient heuristic. Second, Klahr and Dunbar (1988) employing a generation paradigm in a scientific inquiry task found that their subjects could be characterized as being either "theorists" or "experimenters," depending upon their preferred mode of inquiry.

A number of studies have been carried out to assess developmental trends in scientific abilities. Moshman (1979) investigated the development of formal hypothesis testing ability in a selection paradigm. Even amongst Moshman's oldest sample (college students) only a third regularly acknowledged the use of falsification. Moshman implicates an inability to distinguish between hypotheses (theory) and evidence (data) in explaining this lack of experimental sophistication. This is precisely the conclusion drawn by Kuhn and her associates (Kuhn, Amsel and O'Loughlin, 1988; Kuhn, 1989), in explaining the differences between children and adults as scientists. A recent study by Sodian, Zaitchik, and Carey (1991) indicates that the age differences reported above might apply only in a task which demands the generation of a critical hypotheses.

A recurrent finding in studies of rule discovery is the less than optimal performance by subjects. Perhaps this finding reflects the effects of employing an unsuitable strategy at a particular stage of the inquiry process. That is, it is possible that subjects were forced to employ an inappropriate mode of inquiry either because they were forced to continue with a less efficient method beyond the point at which they would have employed a more powerful one, or because they were provided with a powerful method when not yet ready to extract the maximum benefit from it.

It follows that a more successful paradigm for the study of inquiry would be to provide subjects with an

environment in which they have the opportunity to choose between data-driven and theory-driven modes of inquiry - between observation alone and active experimentation. One may then observe whether one of the methods is preferred over the other at some or all stages of the inquiry process. Will any characteristic patterns of transition between modes of inquiry emerge? Will we find a more efficient use of experimentation? If groups of subjects differ in their respective use of either or both modes of inquiry how do they differ? How, if at all, would the incidence of positive instances in the inquiry environment affect subjects' choice of method? Finally, will there be any developmental trends in the use of the two modes of inquiry?

Method

Subjects. The subjects were 27 college students and 27 8th graders.

Procedure and Materials. Each subject solved three problems. Items were created through combinations of six binary dimensions. The subject's task was to distinguish between members and non-members of a target group.

A computer-based environment allowed the subject to freely choose between the reception or generation mode of inquiry. On each trial the subject could request a ready-made item or construct one (by choosing the desired values from a menu). In either mode, the subject had to state his/her belief regarding the categorical membership of the item and received feedback on the accuracy of that belief. In addition, the subject could choose a "Test Me" option on any trial (referred to as the Exit Sequence) in which s/he had to correctly classify ten consecutive test items in order to prove that s/he had learned the rule. An error during the Exit Sequence returned the subject to the main menu.

Design. The experiment had a three-way factorial design with factors being percent of positive instances supplied by the environment (10%, 25%, or 50%), number of Critical Values (1,2 or 3), and Problem Isomorph (one involved a personnel director evaluating job applicants,

another involved trees which grew or did not grow in a particular type of soil, and a third involved a collection of geometric shapes). Each subject completed three problems, organized according to a Greco-Latin Square design.

Results and Discussion

Each step in the subjects' protocol was classified as one of four behavior types: "Get" (a request for ready made item), a "Positive Build" or a "Negative Build" (the generation of a test item by the subject whereby "Positive" and "Negative" refer to whether the subject believed the item built to be a positive or negative instance of the target group), or an "Exit" (a request to enter the Exit Sequence).

A preliminary analysis established that problem content did not significantly affect any one of nine basic measures. Consequently, content does not serve as a factor in the analyses.

Global Measures of Inquiry Behavior

To characterize the overall nature of the inquiry behavior we tabulated the frequency of the four major behavior types. The results, including a comparison between age groups, are presented in Table 1.

Adults used the generative mode more often than did children. They used it on 29.1% of the trials (5.5% for the children), employed it at least once in 79% of the problems (35% for children) and 25 out of 27 adult subjects used it on at least one problem (14 out of 27 children). Children relied more on the reception strategy, entered into (and thus, failed out of) the Exit Sequence more often, and took longer to solve problems.

Transitions Between Modes of Inquiry

It was hypothesized that when given the freedom to choose between modes of inquiry, subjects would begin by opting to view ready made exemplars and only later progress to a strategy in which they would generate test items. The mean trial numbers of the four behavior types appear in Table 2.

Measure	8th Graders		Adults		t value	p
	N of Trials	% of Trials	N of Trials	% of Trials		
GET	15.98	62.6	9.18	55.2	3.19	.002
POSITIVE BUILD	1.17	4.6	3.59	21.6	-4.95	<.001
NEGATIVE BUILD	0.23	0.9	1.25	7.5	-3.76	<.001
EXIT	8.15	31.9	2.61	15.7	3.94	<.001
TRIALS	25.53	100.0	16.63	100.0	3.11	.002

Table 1: Means and t-Values for Behavior Types by Age

Behavior	8th graders ^a	Adults
GET	8.04	6.54
POSITIVE BUILD	11.37	8.97
NEGATIVE BUILD	12.89	10.52
EXIT	14.60	14.32

Table 2: Mean Trial Number for Four Categories of Inquiry Behavior

^a 8th grade statistics reflect only those protocols in which the subject performed at least one "Get" step and one "Build" step.

The mean trial numbers were in the predicted direction both for the adults and for the 8th graders who used the generative procedure. Moreover, the results indicate that within generation the more common positive test precedes the less frequently used negative test.

We then checked whether or not the four behaviors come in clusters (i.e., if each inquiry behavior was most likely to have been preceded by a behavior of the same type). As it turned out for both age groups, every behavior type was most likely to have been preceded by a like behavior. This indicates that beyond the tendency to first observe and later experiment, subjects progressed through the process in an orderly fashion in which once a stage of inquiry was entered they tended to stay in it for a while before going onto the next stage.

Our next question concerned the effect of mistakes on the progression through the inquiry process. We compared the number of times that a correctly identified generation and an incorrectly identified generation preceded a request for seeing a ready-made item (out of all generations). The comparison revealed that for both age groups there were significantly more instances in which a request for a ready-made item was preceded by an incorrectly identified generation ($t(159) = 1.87$, $p < .05$). Thus, it may be seen that the orderly nature of the inquiry process was preserved in a regressive as well as progressive manner, when subjects encountered difficulties in their investigations.

Environmental Mediation: The Availability of Positive Instances

People engaged in a rule discovery task exhibit a strong confirmation bias, manifested in the tendency to generate positive, rather than negative test items. Our results (see Table 1) show a similar trend, with positive tests outnumbering negative ones by a ratio of 3:1 among adults and 5:1 among the 8th graders.

As it turned out, the tendency to prefer positive over negative test items was mediated by the environment. An analysis of the effect of the percent of positive instances in the environment on the total incidence of positive tests revealed that the number of positive tests increased as the percentage of positive instances in the environment decreased ($F(2,154) = 4.76$, $p = .01$). Furthermore, the analysis revealed a significant interaction between age and percent of positive instances ($F(2,154) = 3.17$, $p < .05$), with the above trend observed only amongst the adult sample.

Characteristics of Experimentation

Experimentation calls for the controlled manipulation of variables in order to study their effects. To gain an insight into the nature of experimentation in the present task, we recorded the number of values changed in each generated item relative to the item just preceding it. We then checked which independent variables were related to the number of changes. It was found that fewer values were changed when: 1- the

previous item belonged to the target group than when it did not belong to it (1.84 vs. 2.66 changes; $F(1,433) = 39.03$, $p < .001$); and 2- the previous trial already involved a generation rather than a reception or a failed attempt at the Exit Sequence (2.04 changes following a Positive Build, 2.19 changes following a Negative Build and 2.82 following a Get or Exit; $F(2,433) = 18.61$, $p < .001$).

Interestingly, in this analysis age was not found to be a critical factor. Apparently on the occasions that children did construct an item, the strategies they adopted were similar to those employed by the adult builders.

Individual Differences

For a final analysis, we analyzed the profiles of subjects' behavior in search of consistent individual differences in the process of inquiry. In view of the strong effect of age, we carried out the analyses separately in each age group.

For the adults, the factor analysis revealed two main factors which included 22 of the 27 subjects (10 subjects in Factor I, 12 in Factor II). Adult Factor I subjects experimented earlier as well as more frequently than did their Adult Factor II counterparts. In contrast, Adult Factor II subjects were characterized by a more frequent use of the reception mode of inquiry. Factor I subjects were also faster to solve the problem (14.7 trials per problem vs. 18.8 for the others).

Among the 8th graders there was one prominent factor including 17 of the 27 children. These children were quite similar to Factor II adults, exhibiting strong preference for the reception rather than the generation mode of inquiry. Two additional types of behavior were observed amongst the children, consisting of 4 and 5 subjects, respectively. Both types were characterized by a tendency to follow a correct response with the initiation of the Exit Sequence. 8th graders included in Factor III solved their problems in fewer steps than did the 8th graders in Factors I or II.

In summary, our results demonstrate a complex but coherent inter-relationship between the reception and generation paradigms.

This inter-relationship is marked by a clear progression in the transition from one to the other. In addition to consistent individual differences in the use of these strategies, their use is related to age and the availability of positive instances in the environment. We propose that any future study of rule discovery should provide subjects with the ability to use both the reception and the generation modes, and to freely switch between them. Such a paradigm, and the study of variables which affect transitions between modes of inquiry, would greatly enhance the future study of rule discovery behavior.

References

- Bruner, J., Goodnow, J. and Austin, G. 1956. *A study of thinking*. N.Y.: Wiley.
- Evans, J.St.B.T. 1989. *Bias in human reasoning: Causes and consequences*. London: Erlbaum.
- Klahr, D. and Dunbar, K. 1988. Dual space search during scientific reasoning. *Cognitive Science* 12:1-48.
- Klayman, J. and Ha, Y. 1987. Confirmation, disconfirmation and information in hypothesis testing. *Psychological Review* 94:211-228.
- Klayman, J. and Ha, Y. 1989. Hypothesis testing in rule discovery: Strategy, structure and content. *Journal of Experimental Psychology: Learning, Memory and Cognition* 15:596-604.
- Kuhn, D., Amsel, E. and O'Loughlin, M. 1988. *The development of scientific thinking skills*. Orlando, Fla.: Academic Press.
- Kuhn, D. 1989. Children and adults as intuitive scientists. *Psychological Review* 96:674-689.
- Moshman, D. 1979. Development of formal hypothesis-testing ability. *Developmental Psychology* 15:104-112.
- Mynatt, C.R., Doherty, M.E. and Tweney, R.D. 1978. Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental*

Psychology 30:395-406.

Sodian, B., Zaitchik, D. and Carey, S. 1991. Young children's differentiation of hypothetical beliefs from evidence. *Child Development* 62:753-766.

Wason, P.C. 1960. On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12:129-140.

Wason, P.C. and Johnson-Laird, P.N. 1972. *Psychology of reasoning: Structure and content*. Harvard University Press.