AN INVESTIGATION OF PROBLEM SOLVING STRATEGIES USED BY STUDENTS IN SOLVING BIOLOGICAL CLASSIFICATION TASKS: PILOT STUDY

Jeffrey W. Bloom
February, 1986

The present study focuses on problem solving in science education. Students are rarely asked to solve poorly structured problems, such as those encountered in real life situations (Frederiksen, 1984) or problems that require them to draw on information from several units of study. When presented with problems that lack adequate structure in terms of completeness of givens and available solution strategies and that require knowledge from several different units of study, how do students go about solving the problems? What strategies or heuristics are used? In what ways do students use their prior knowledge? How does prior knowledge affect the problem solving process? Are certain strategies more effective at solving problems in terms of efficiency and utilization of appropriate prior knowledge?

The following section discusses the structure of knowledge and how it relates to problem solving in the domain of science. This is followed by a discussion of the theoretical background of cognitive science and problem solving. The methodology for obtaining verbal protocols of the problem solving process is discussed briefly in the next section.

Background

Discipline of Science

The structure of knowledge in science is characterized by a framework of theories and concepts. It is this framework of knowledge that guides the processes of scientific inquiry (Schwab, 1964). As Novak (1984) describes, "what scientists do is to construct explanatory models and this process is continuous, with each new construction dependent on the current 'population" of ideas — all of which are undergoing gradual change" (p- 2).

The knowledge used to construct models comes out of the methodology of science or scientific inquiry. Schwab (1964) has differentiated two types of inquiry, stable and fluid inquiry-stable inquiry is concerned with specific pieces of larger theoretical or conceptual frameworks, the body of scientific knowledge- The methods of science are used to construct and reconstruct the body of knowledge. In stable inquiry, specific problems are tackled in an attempt to better understand how a particular phenomenon functions or relates to other phenomena- In essence, according to Schwab, "...science is a process of <u>constructing</u> bodies of <u>tentative</u> knowledge, of discovering <u>different</u> ways of making data coherent, and 'telling' about a given subject matter" (p. 35).

Fluid inquiry, on the other hand, takes a broader view of the discipline- The primary concern is with the validity and coherence of the body of knowledge and how scientific knowledge fits in with the knowledge of other fields.

In both stable and fluid inquiry scientists involve themselves in different ways of thinking and approaching problems. This is most aptly put by Bateson (1972), "As I see it, the advances in scientific thought come from a <u>combination of loose and strict thinking,</u> and this combination is the most precious tool of science" (p. 75). Thinking, in whatever form, is a tool for perceiving and solving the problems of concern to science.

In looking at science, the body of knowledge and the methodology become a web of interrelatedness. The knowledge arises out of the methodology and the methodology draws on the body of knowledge in order to progress- With problem solving as the central theme of the methodology, problem solving is viewed as being domain specific- It is domain specific because it requires specialized information from the body of scientific knowledge.

Coqnitive Psycholoqy

At the core of the information—processing theory is the distinction between long-term (LTM) and short-term or working memory (WM). Long-term memory is a more or less permanent store of information or declarative knowledge and processes or procedural knowledge. The basic unit of declarative knowledge is referred to as a proposition. Each proposition is linked or related to others, resulting in what is hypothesized as a propositional network. Procedural knowledge is represented as productions. Productions are specific processes or rules which actively transform information. Productions operate rapidly and automatically, whereas the activation of propositional networks tends to be more conscious and significantly slower. It is thought that productions are stored along with associated propositions throughout propositional networks in LTM. In this way when a specific process or production is needed it is readily available due to its proximity to related information (Gagne, 1985).

As propositions and productions are activated they are, in a sense, placed in working memory. All newly received information is temporarily stored in WM, whether that information comes in through the senses or from LTM. It is in WM that ideas are manipulated and related to other ideas (Gagne, 1985). Active portions of propositional networks are stored in WM along with other propositions in the process of being constructed. The critical limitation of WM is its limited capacity. Only five to seven items of information or propositions can be held in WM at any one time (Gagne, 1985; Frederiksen, 1984). Frederiksen notes that the capacity of WM is a critical factor in the size of problems that can be dealt with at any given time. He summarizes the contention of other researchers (Battig & Bellizza, 1979; Miller, 1956; Tulving, 1962) that one way to increase the capacity is through "chunking."

> Complex cognitive structures can be developed that allow a single symbol or concept to represent a collection of related items of information....Chunking helps make it possible to process a great deal of detailed information automatically… (Frederiksen, 1984, p. 365)

Problem Solving

Most research and development in problem solving has either been generic (focused on general problem solving skills) or based in physics, mathematics, or chemistry. Problems in these areas tend to be characterized by one right answer and a definite sequence of problem solving steps. Frederiksen (1984) describes the situation in schools:

Schools seldom require students to solve… fuzzy problems — problems that are not clearly stated, where the needed information is not all available, there is no algorithm, and there may not be a single answer that can be demonstrated to be correct (p. 363).

He cites two categories of problem types from Simon (1978): "well-structured problems" and "i11-structured problems." Well-structured problems are defined as, (a) being clearly formulated, (b) having a known algorithm, (c) having specific criteria for evaluating the correctness of the solution, and (d) having one correct answer. Ill-structured problems, on the other hand, are defined as, (a) not being clearly formulated, (b) lacking procedures that guarantee a solution, (c) lacking specific criteria for evaluating the correctness of the solution, and (d) possibly having more than on correct answer. In addition, Ill-structured problems tend to be more complex. It is frequently difficult to know when a solution has been attained. Frederiksen (1984) points out that there are, however, no clear cut division between these two types of problems. Problems tend to fall along a continuum. He has labeled those problems that lie between the two extremes as "structured problems requiring productive thinking." In general, these problems are similar to well-structured problems, but require the problem solver to generate certain procedures. Most problems presented to students in school tend to be well-structured. In biology laboratories students have explicit steps to follow and one correct answer to find. Most of the research on problem solving has involved well—structured problems, as well (Greeno, 1980).

Two basic distinctions in problem solving procedures are described as algorithmic and heuristic (Cyert, 1980). Algorithmic procedures are commonly associated with mathematical and computer programming problems. The sequence of steps is highly structured and relatively inflexible. Heuristic procedures, on the other hand, tend to be more global in their application, because of the flexibility inherent in the way in which they can be applied by the individual.

Two basic distinctions in problem solving procedures are described as algorithmic and heuristic (Cyert, 1980). Algorithmic procedures are commonly associated with mathematical and computer programming problems. The sequence of steps is highly structured and relatively inflexible. Heuristic procedures, on the other hand, tend to be more global in their application, because of the flexibility inherent in the way in which they can be applied by the individual.

> Heuristic procedures, in contrast to algorithmic (deductive) ones, don't guarantee anything. They give answers which are good enough most of the time. The errors they produce tend to be systematic rather than consequence of...heuristics is that a path is opened up to consider problems of theory structure and the growth of knowledge jointly. (Griesemer, 1985, p. 214)

Examples of heuristic procedures include: (a) means-end analysis, (b) hypothesize-and-test, (c) use of analogs (replacement of original problem with an abstracted version), (d) dividing the original problem into sub-problems (Frederiksen, 1984), and (e) a variety of domain-specific heuristics (Scandura, 1977).

Other distinctions in the methods of problem solving i nclu.de,

> ...weak <u>methods</u> — problem solving techniques of quite general application whose generality is assured by the fact that they do not use or require much prior knowledge of the structure of the problem domain.... [and] <u>strong methods</u> —

> powerful techniques that are carefully tailored to the specific structure of the domain to which they are applied. (Simon, Langley, & Bradshaw, 1981, p. 5)

The authors classify generate-and-test (or hypothesize-and-test), means-end analysis, and heuristic search (process of modifying previously tried solution possibilities) as examples of weak methods of problem solving. These methods tend to be characteristic of novice problem solvers, however, they are useful to experienced problem solvers when working with novel situations or new knowledge domains. Related to the weak/strong distinction is the novice/expert dichotomy (Larkin, 1980). The primary distinction between these two is that experts tend to have a larger and more organized knowledge structure in a specific domain. In other words, experts have a more coherent organization of the information and processes of their particular domain. The problems encountered by experts tend to be more complex. Reif (1980) explains that,

> ...as problems become more complex and require more knowledge, the selective retrieval of information becomes exponentially more difficult. Hence the solution of complex problems, unlike that of simpler problems, depends crucially on efficient search procedures and forms of knowledge organization.(p. 39 - 40)

As Greeno (1980) paints out, "all problem solving is based on knowledge" (p 10). The knowledge or information available to the problem solver plays an essential role in the solving of problems. Novak (1984) adds that both the knowledge structure of science and cognitive theory play an important role in science education. In fact, he notes that cognitive science "...places central emphasis on the role that concepts and conceptual framework play in human construction of meaning" (p. 1). Many other researchers are coming to the same conclusions and are adopting the primary theory of cognitive science, the information processing theory, as a framework in which to base their research efforts.

The efficiency of cognitive functioning (information recall, problem solving, concept development, and so forth) is highly dependent upon the organization of declarative knowledge or the way in which propositions are related or connected in the propositional network. In other words, prior knowledge plays a significant role in effective and efficient thinking. Gagne (1985) states that the, "learning of declarative knowledge is synonymous with the creation of meaning. When no meaning (no connections) can be created, nothing is learned" (p. 79). Bateson (1979) referred to the gaining of meaning as the process of relating a complex or knot of connectedness or connected ideas (called relevance) to the context in which it is seen. Similarly, Ausubel (1963) emphasized the role of prior knowledge in meaningful learning. In science education, Champagne and Klopfer (1984) contend that, "the new cognitive perspective directs greater attention to the structure of the student's knowledge and to the influence of science—specific knowledge on the student's acquisition of science information and concepts" (p. 93). It is worth noting here that a student's prior knowledge may include misinformation. If a student developed misconcepts at an earlier time, these misconceptions are connected, somewhat tenaciously, to other propositions in the propositional networks in LTM. If, as Champagne and Klopfer suggest, "...students' comprehension of science instruction is greatly influenced by their existing knowledge" (p. 101), then any misconceptions they have may have tremendous impact on what and how they learn, as well as how they go about salving problems.

In light of the information processing theory, Shulman and Carey (1984) refer to the individual learner as "boundedly rational" due to the constraints of their own information processing capacities. They continue that, "because we lack the cognitive capacities to apprehend the world as it is, we are forced to construct representations of that world and to engage in thinking and reasoning within the confines of those constructions" (p. 508). In other words, these representations are simplifications of what we perceive. The processes that are responsible for developing these representations are based on prior knowledge or the structure of the propositional networks and embedded productions. In the language of problem solving, these representations are referred to as "problem spaces." The "task environment," on the other hand, is the perceived situation which is then translated into the problem space (Newell & Simon, 1972; Frederiksen, 1984; Shulman & Carey, 1984).

Summary

In order to develop a working knowledge of problem solving in science education, we need to consider (a) the structure of the knowledge of the discipline, (b) in what form students remember or store the information that is presented to them, (c) the structure of problems, (d) the various heuristics of problem solving, and (e) how students process the information in problem solving tasks. The relations and interplay between these aspects comprise the context of student problem solving in science. For instance, in considering a problem that is not well-structured, but is in the realm of structured problems requiring productive thinking, we can expect certain items of information from the knowledge base of the discipline to be included in the problem statement. It is then expected that the student will make a connection between the given information and related information in his or her memory. A representation of the problem will then be formed in WM, which will consist of propositions from LTM and new information from the task environment (statement of the problem). At this point, various productions (automatic processes) will manipulate and transform the information in WM. If productions are unavailable, the student may have various strategies stored as declarative knowledge, which could be brought into WM and followed step by step in a slower and more conscious manner. Upon solving the problem, the solution would result in one or more new propositions which would then be added to the propositional network in the vicinity of the propositions used to help solve the problem.

Methodology

The investigation of how students think during problem solving tasks requires a methodology that exposes the thoughts involved in attempting to solve problems. The think aloud technique of obtaining verbal protocols is one way of tapping into the thought processes of students who are actively engaged in problem solving. The think aloud technique has two major forms: (a) concurrent verbalization and (b) retrospective verbalization (Ericsson & Simon, 1984). Both forms are effective tools for obtaining verbal protocols of student thinking. In the following section, the theoretical basis for the think aloud technique and protocol analysis will be discussed.

Theoretical Framework

Verbal records have been used as data for decades. With the advent of behaviorism verbal data was considered to be less than legitimate. However, with the technological advances, such as tape recorders, and with the falling away of behaviorism as the central theory in psychology, the use of verbal data has been increasing in popularity (Ericsson & Simon, 1984; Larkin & Rainard, 1984; Hayes, 1981).

Verbal data are referred to as verbal protocols. The term "protocol" is a descriptive sequence of activities that occur during the performance of a task (Hayes, 1981). Protocols can involve any number of activities, nonverbal as well as verbal. In the following discussion, however, the focus will be on verbal protocols.

The collection of verbal protocols generally involves asking students to say out loud whatever they are thinking, whether it makes sense or not. The responses are tape recorded, then is either declarative knowledge or newly perceived or constructed information (Ericsson & Simon, 1984). Procedural knowledge, because of its characteristic automaticity, is not vocalized. So, vocalizations will only contain (a) information that is retrieved from LTM in the form of declarative knowledge, (b) information that is newly perceived, or (c) information that is being generated or constructed in WM (productive thinking).

With a verbal record in hand, it is up to the researcher to make inferences as to how and why certain events (sequence of thoughts) take place. Fortunately, there are certain landmarks that can help the researcher cue into points at which nonverbalized events are taking place and into the nature of the information. As mentioned in assumption six, pauses and hesitations indicate shifts in the processing of information. For example, a pause in the flow of thinking may indicate the retrieval of relevant information from LTIi or the manipulation of information in WM. The retrieval or manipulative processes most likely contain activated procedural knowledge which is operating automatically and without the student's conscious control. Such processes are not vocalized. A sequence of coherent thoughts with very few pauses is indicative of information that has been retrieved from LTM. This information tends to be highly organized in memory and is merely recalled and vocalized by the student. On the other hand, a sequence of coherent vocalized thoughts embedded with many pauses probably indicates a set of information that is being reconstructed and processed into new propositions in WM (Ericsson & Simon, 1984). Although bits (and possibly all) of the information may have come from LTM, the transcribed verbatim. The resulting protocols are then analyzed using any of a variety of methods.

In the -following discussion o-f verbal behavior, it is important to distinguish between the terms: verbalized and vocalized. "Verbalized" refers to thoughts in verbal form, but which are not necessarily spoken aloud or vocalized. Of course, if the discussion refers to verbal data or verbal protocols, it can be assumed that vocalizations are the source of the verbal i nf or mat i on.

Ericsson and Simon's (1984) model of protocol analysis serves as the basis for the methodology of this study. It is therefore necessary to consider some of the theoretical assumptions and implications of protocol analysis and the thinking aloud technique.

The use of thinking aloud verbalizations as data is based on several assumptions as to the nature and meaning of verbal behavior. (1) Any thought that is verbalized corresponds to information in working memory (WM) . A thought that is verbalized is the information on which a student's attention is focused. (2) The vocalized information represents information in WM that has been verbally encoded. The vocalized thought is equivalent to a thought in verbal form in WM. If a thought is not in a verbal form in WM then the vocalized form has resulted from a verbal encoding or transformation of the original thought. (3) Verbalizing begins when a thought is noticed or heeded. The verbalizations may not be vocalized. However, verbal processes begin the moment a thought is heeded. (4) The structure of the thought in the focus of attention is reflected in its verbalized form. For nonverbal thoughts, the structure is reflected in its verbally encoded form. (5) The way in which thoughts are articulated will correspond to the organization of knowledge in LTM. Specifically, an articulated unit (phrase, sentence, etc.) vocalized without hesitation will reflect the organization of the propositional network being accessed or activated. (6) Pauses and hesitations in the vocalization of thoughts can be used to signify shifts in the processing of information (Ericsson &

Simon, 1984). For example, a pause may indicate the retrieval of more information from LTM or the manipulation of information in WM.. These six assumptions represent the working basis for the use of verbal protocols as data. In essence, the information articulated is the information on which attention is focused. The acquired verbal data depict the structure of propositional knowledge and provide indicators of processing shifts from which inferences about the processes can be made.

The transcript generated from a thinking aloud session does not represent a complete record of all the students thoughts. However, it is a systematic record of the student's thinking over time (Larkin & Rainard, 1984). Reading a transcript is somewhat analogous to watching a sports event on television (without listening to the commentators). The viewer does not perceive every event, but does see a certain proportion of the events including the highlights. The result is a fairly accurate picture of the sequence of events over time.

It is important to recognize that the information vocalized organization of the information is new.

The gathering of verbal data can be accomplished in two major ways, concurrently or retrospectively. Concurrent verbal reports require the student to think aloud as thoughts arise during a task. Retrospective reports are obtained by asking the student to remember what thoughts occurred in a task that was just completed. Both of these reporting techniques have a number of variations which can be used to meet the specific needs of the researcher, however only those that apply to the present study will be included in the following discussion.

Concurrent verbal reports have the advantage of accurately showing what thoughts were heeded during the task and of representing an accurate sequence of the thinking involved in a task. The thoughts that arise are the thoughts that are being heeded at the moment. The sequence of thoughts determined concurrently is obviously the sequence that actually occurs. Retrospective reports, on the other hand, may include thoughts that actually occurred during the task. However, it is likely that the reporting of these thoughts is embellished with new and related thoughts that arise at the time of the reporting. The memory of the actual sequence of thoughts in retrospective reports may not be accurate. However, the use of retrospective reports as a follow-up to concurrent reporting can be very helpful. Since not all thoughts are vocalized in concurrent reports, retrospective reports have the advantage of being able to fill in the gaps (Ericson & Simon, 1934; Larkin & Rainard, 1934; Hayes, 1981).

Establishing the concurrent reporting session is relatively straight forward- First, instructions are read to the student that simply ask him/her to say whatever he/she is thinking, whether it makes sense or not. Then a practice problem is given to the student, after which the student and researcher can discuss the think aloud technique (Ericsson & Simon, 1984; Larkin & Rainard, 1984). As a general rule the researcher says as little as possible during the session, thereby avoiding influence upon the student. Exceptions are predetermined and listed for reference by the researcher. These exceptions include reminders to talk aloud or talk louder, nondescript comments (okay, hmm), and technical comments and instructions. The basic idea is to interfere as little as possible, but to maintain a certain amount of control to ensure adequate data collection.

Working with a retrospective reporting session is more tricky. How this session is conducted is dependent upon what is expected from the data. If what is desired from the student's report is to fill in gaps in a concurrent report, then asking for a general summary of what the student was thinking during the task is the best first step. The student is not being influenced by leading questions. The response will tend to be a more accurate depiction of what occurred. However, new information may be brought into the response, but will tend to be closer to what was active during the task. A more general summarizing question forces the student to rely on the contents of WM directly or on pointers in LTM. In other words, if the information is no longer in WM it has either dropped out or has been stored in LTM. The information that remains active in WM can act as a pointer that re-activates the information in LTM.

A more directed or specific question may yield the type of information desired by the researcher, but may influence the response by the student. Instead of relying on what can be remembered about the task, the student may answer the question without relying on a memory of the task. For example, if the student is asked why he/she came up with a specific answer. The response may be based on a motivation to justify the answer with whatever information can be generated even though the latter information may not have been accessed during the task. This is not useless information for the researcher. It may provide some sense of the student's extent and organization of knowledge of the answer, but it does not reflect the thinking that took place during the task.

Once the verbal reports are tape recorded the tapes must be transcribed. Since pauses, intonations, and seemingly extraneous comments may be important cues to nonverbal processes, the transcribing process must include ways of displaying pauses and various grunts and groans. Larkin and Rainard (1984) recommend separating each segment (vocal sequence without any pauses) by a blank line. No punctuation is used in the transcripts. Intonations and other characteristics are included in bracketed comments. Completed, the transcripts are then ready for analysis.

The way in which the transcribed protocols are analyzed varies with the goals of the study. Ericsson and Simon (1984) recommend that a theoretical model be developed before collecting data. They do admit that this is not absolutely necessary. Larkin & Rainard (1984), on the other hand, do not view having a predetermined model as necessary at all. They do, however, have the development o-f a model as the primary goal o-f the study. Since the data collection methodology strives not to interfere with the student's thinking processes, the development of a model does not necessarily need to precede data collection. Having fewer preconceptions about the results may, in fact, allow the researcher to take a fresher and more objective view of the data.

The first step in the analysis process is usually to develop a scheme for segmenting the transcript. If the transcripts are typed as described above with no punctuation and blank lines dividing vocal sequences then a detailed segmenting process is achieved. However, the task may have characteristics that allow the researcher to segment the transcripts into larger chunks. In the present study, it was found that larger segments helped make the analysis process more manageable. The task the students were involved in had givens (i.e., given information) presented at various points throughout the task. Each given initiated a new sequence of thoughts and therefore presented an obvious cut off point for a new segment. Each segment, in whatever way it is created, can be defined by the researcher as a working unit or units prior to the next stage of the analysis process.

At this point, the researcher needs to make a number of decisions, if not already made, about how to proceed. The data, in raw form, are rich in possibilities for analysis. The richness itself can be somewhat overwhelming. In essence, the researcher needs to determine what variables are present in the data and to decide on which ones to consider. Once delineated the variables can be reduced to simpler form by means of an encoding process. For example, if the variables are steps in an algorithm or heuristic of problem solving then the segments or words that exemplify these steps can be reduced to code words for each step. Ericsson and Simon (1984) recommend that the number of code words or variables be kept to a minimum. Ten to fifteen variables are optimal for a manageable analysis. Of course, the transcripts can be analyzed more than once to look for different aspects of research questions.

several theoretical assumptions underlie the process of interpreting and encoding the transcribed data. Ericsson and Simon (1984) have delineated four such assumptions that incorporate contemporary theories of problem solving in light of information processing psychology. (a) A student's involvement in a problem task is viewed as a search of a problem space. The process of searching includes an accumulation of information or knowledge, which is not necessarily correct. (b) Each step in the search process involves a procedural element that operates on the information in WM. The total number of relevant procedures available is relatively small. The activation of each procedure changes the nature of the problem space. (c) The vocalizations of the student are equivalent to some part of the information that is being held in WM. (d) The information in WM is mostly in the form of knowledge that is needed for the operation of procedures, new knowledge produced by the procedures, and various goals and subgoals.

After the encoding process the resulting protocol is analyzed in terms of the theoretical assumptions just discussed and the theoretical framework of the research problem under investigation. The encoded protocols are a skeletal sequence of the events that take place within the task environment. It is the job of the researcher to infer a more complete picture of what takes place (Larkin & Rainard, 1984). For example, the state or condition of the information in the problem space may be encoded at one point. Following a pause a new condition is defined by new information. The action or procedures involved in this change of state are inferred by the researcher. The nature of the change may create a rather definite outline of what is required of the action. How the action is defined and labeled depends on the theoretical framework guiding the study. Smith (1971, cited in Ericsson & Simon, 1984), for instance, used Guilford's Structure of the Intellect as a model for defining the contents, operations, and products of his encoded protocols.

Task Environment

The task environment in the present study consists of a computer-based biology problem solving game. Animal Tracks (Bloom, 1985a). The goal of the game is to determine the kind of animal described by a series of clues. Six possible clues are selected by the student from the following categories: food, air, water, reproduction, protection, and ecology/behavior. The clues are based on a theoretical framework of evolution and ecology. Each clue described some feature of how the animal in question is adapted for survival in terms of the framework of the category. The information contained in the clues consists of one or two propositions. After each clue appears the student has the opportunity to pass and select another clue or guess the animal. The information in the first five clues provide a definitive description of the animal in question.

Most of the information contained in the clues is covered in a typical high school biology text. Out of necessity for accuracy some information is beyond the scope of high school biology. The remaining clues in such a situation are sufficient for defining the animal in question.

As the student proceeds, each clue provides an incomplete set of givens. The structural nature of the problem changes as the student selects each clue. With one set of givens appearing in each clue the structure becomes increasingly structured as the student proceeds through the clues. The degree of structured-ness, however, depends on the prior knowledge of the individual student. If the student does not understand the information in the clue, then that particular given cannot be included as a factor in defining structure of the problem. In essence, the task environment presented to students in this study represents a variable degree of structured-ness. The problems presented to the student fall within the category of structured problems requiring productive thinking as defined by Frederiksen (1934).

Each student completed from three to six trials (animals in Animal Tracks). The order of clue categories was selected by the student.

Subjects

The subjects were 10 high school students enrolled in an advanced biology course in a public school in the Houston metropolitan area. Students were selected randomly — from the total number who agreed to participate and had received parental permission.

Procedures

Background In-formation. Prior to the study, the students were given a number to be used for identification in the study. Each student was then requested to fill out a brief questionnaire, which included: gender, age, ethnic background, and grades in previous science courses. The Greqorc Stvle Delineator (Gregorc, 1982) was then administered as a measure of student learning style.

Concurrent Verbal Reports. The verbal report session began with a short conversation with the student. The purpose was to help the student relax. Following this the researcher read aloud the instructions to the thinking aloud session (see Appendix A), which was followed by a practice problem. The practice problem gave each student the opportunity to think aloud and get feedback on their performance of the methodology. Then the student was given the opportunity to read the instructions to Animal Tracks and ask questions.

All vocalizations from the practice problem and the actual problem solving session were tape recorded. Each student was asked to begin Animal Tracks. A list of accepted comments (see Appendix B> was used as a guide by the researcher to govern his comments during the session. The researcher interfered as little as possible. At the completion of each animal the researcher requested the student to wait and respond to several questions (retrospective reporting).

Retrospective Verbal Reports . Following each trial each student was asked to summarize his/her thinking during the trial and/or to answer specific questions about what his/her thinking at specific points in the trial.

Results

Data Analvsis

The tape recordings were transcribed verbatim according to the guidelines established by Ericsson and Simon (1984) and Larkin (1984). The points where clues were accessed were inserted in brackets along with other comments of the researcher.

Analysis of the transcripts reveals four general categories of directed thinking and several subdivisions of each, as follows:

(1) repeating or rephrasing of givens (clues)

   (a) correctly repeating or rephrasing given

   (b) incorrectly rephrasing given

   (c) total of all repeats and rephrases

(2) conjectures or hypotheses

   (a) positive conjecture of possible answer

   (b) eliminating a possible answer

   (c) total conjectures

(3) testing conjecture against givens

   (a) verifying conjecture against givens

   (b) refuting conjecture from givens

   (c) weak test of conjecture against givens

   (d) total tests of conjectures

(4) other productive thinking (inferences)

   (a) proximal inferences (includes classifying animal in question in a taxonomic group)

   (b) distal inferences (includes relating givens or constructed information to (ecological or behavioral) context and describing or inferring characteristics of animal in question)

   (c) misconceived inferences

   (d) total inferences

The frequency of each of item was calculated for each student across all trials. In addition, the frequency of each item was calculated for successful trials (correct answer was obtained) and unsuccessful trials (correct answer was not obtained). A total of all directed was determined, along with totals for successful and unsuccessful trials. Student scores in each category were calculated by dividing the frequency of the item by the number of clues (givens) used by the student. A degree of success score was calculated by dividing the number correct by the total.

The above items were also used to construct a decision tree for each trial- The decision trees were analyzed for patterns of effective and ineffective strategies and for other descriptive observations. Examples of a few characteristic decision trees appear in Figures 1 through 5.

A correlation matrix of the combined (successful and unsuccessful) scores of all items appears in Figure 6. Correlations of between the degree of success score and the totals of the combined, successful, and unsuccessful scores in each category appear in Figure 7. Mean and standard deviations of the totals of combined, successful, and unsuccessful scores appear graphically in Figure 8.

Discussion

The most common heuristic evident in the data is a process of generate (a possible solution or conjecture) and test. The frequency of conjectures generated per clue are the single most dominant form of directed thinking, followed by the repetition of givens and the testing of conjectures against the givens (see Figure 8). These three categories of directed thinking comprise the majority of activity involved in the generate and test heuristic. The role of repetition appears to keep information active in STM and to be useful as a search process in which students engage. The repetition of a given acts to focus *in* attention on the particular proposition being set forth by the given. The major process occurring in conjunction with the repetition of clues is the search of declarative knowledge for propositions related to the proposition of the given. The success of the search process is highly dependent on the extent, accuracy, and organization of prior knowledge.

For example, in Figure 2, the student makes a number of connections to prior knowledge throughout the trial, but the information recalled or generated is either faulty or not well organized. Even though she generates the correct solution in one conjecture, she -fails to ultimately select the correct answer. In Figure 1, the student reads the air clue (skin is moist and lungs of adults are without a diaphragm). He repeats "skin is moist" and "eats meats" (from a previous clue). He then conjectures that maybe it is a "snake," and continues by saying that "snakes have moist skin, well they have slippery skin and they eat meat." The limitation of the extent of knowledge is evident in his choosing to ignore the second proposition of the air clue (which is brought out later in the air clue segment). However, as a whole this student's extent of knowledge is greater than most other students. As can be seen in the Figure, he recalls a great deal of information with which to work. Other students (not included in the figures) recall very little information in connection with the givens. The accuracy of information is faulty in that he relates snakes with having moist skin. Organizationally this student appears to have a large number of clusters of information in memory. He makes many connections in terms of tests of conjectures. However, the lack of efficiency of the whole process seems to indicate a lack of a more comprehensive organization. In contrast the student in Figure 5, verbalizes very little information, but keys in on the animal very quickly. She makes connections to appropriate information efficiently. Both of these students had degree of success scores above the mean (.45), however the score for the student in Figure 1 is .60 correct out of the total number of attempts and the score for the student in Figure 5 is 1.00 (correctly named all animals attempted).

The correlations in Figure 7, support the role o-f the generation of conjectures <R=.85) and repetitions <R=.6i) in successful performance. The testing of conjectures only appear to be significant in unsuccessful trials. In fact, of these three categories of thinking only repetition shows a significant correlation to performance on successful trials, whereas all three categories are significantly correlated with performance on unsuccessful trials. Repetitions appear to be the primary means of focusing attention on the process of searching declarative knowledge. Students working unsuccessfully in a trial, but with higher degree of success overall, engage more actively in these three categories of directed thinking than do students with lower overall degrees of success. The low correlations on successful trials probably indicates that students narrow in on an answer more quickly and are less actively engaged in the generating and testing of conjectures. Further support for the connection between repetitions as a focus of search

strategy and the testing of conjectures is seen in the correlations between these two strategies in Figure 6.

The low correlations and low frequencies of other types of productive thinking (inferences, etc.) to degree of success and other types of thinking are indicative of the lack of use of other strategies for solving the problems. Attempts at classifying animals and describing their structural or ecological characteristics are infrequent and often ineffective when utilized. For example, several students made inferences about the taxonomic group of the animal in question. Frequently this strategy led to no solution at all, or the information generated by the inference was ignored in favor of information in a new clue which led to a faulty solution attempt (see Figure 2, ecology/behavior clue segment). On one occasion a student classified the animal after several wrong, but very good, conjectures. The classification in this instance led to the correct solution (see Figure 3, water clue segment).

The extent, accuracy, and organization of prior knowledge appears to play a significant role in the successful performance of problem salving in the present study. The prevalent heuristic is generate and test. The major components of this heuristic are (a) the generation of conjectures, (b) the repetition of givens as a focusing of attention strategy, and (c) the testing of conjectures against the givens in the clues. As mentioned in the background section, this heuristic is characterized as a weak method of problem solving. Weak methods are commonly used when the knowledge base is not well developed or coherently organized. The low frequencies of inferential and other types of productive thinking indicate the lack of available alternative heuristic procedures. The effective and efficient use of the generate and test heuristic, as with the student in Figure 5, indicates a more highly developed organization of knowledge. This student could be ready to use alternative heuristics, but none were apparent in this study.

The implications of the results of this study point to the lack of a coherent and accurate knowledge-base on the part of a majority of the students. This is a limiting factor in the selection of alternative problem solving strategies and heuristics. If we are to expect students to use higher levels of thinking (inferences, etc.), more emphasis needs to be placed on what information is presented and how it could be presented so that a coherent knowledge-base results. A smaller, but more coherent knowledge-base is ultimately more productive than a larger and more poorly organized knowledge-base.

Further research seems needed in detailing the types of faulty prior knowledge present among students, as well as the further description of the effective and ineffective strategies used by students. What are the characteristics of a coherent knowledge-base as it relates to problem salving in biology? Are there differences between college freshman biology students and high school students advanced biology students? If differences do exist, what accounts for them?

References

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York: Grime & Stratton.

Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballantine.

Bateson, G. (1979). *Mind and nature*. New York: Bantam.

Battig, W. F. & Bellizza, F. S. (1979). Organization and levels of processing. In C. R. Fuff (Ed.), *Memory organization and structure*. New York: Academic.

Bloom, J. W. (1985). *Animal tracks: A biology problem solving game*. Unpublished computer software.

Champagne, A. B. & Klopfer, L. E. (1984). The cognitive perspective in science education. In R. W. Bybee, J. Carlson, & A. J. McCormack (Eds.), *Redesigning science and technology education: 1984 yearbook of the National Science Teachers Association* (pp. 90-103). Washington, DC: National Science Teachers Association.

Cyert, R. M. (1980). Problem solving and educational policy. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.

Ericsson, K. A. & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Frederiksen, N. (1984). Implications of cognitive theory for instruction in problem solving. *Review of Educational Research, 54*(3), 363-407.

Gagne, E. D. (1985). *The cognitive psychology of school 1earning*. Boston: Little, Brown & Co.

Greeno, J. G. (1980). Trends in the theory of knowledge for problem solving. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.

Gregorc, A. (1932). *Style delineator* (Research version). Maynard, MA: Gabriel Systems.

Griesemer, J. (1985). Philosophy of science and "the" scientific method. *American Biology Teacher, 47*(4), 211-215.

Hayes, J. R. (1981). *The complete problem solver.* Philadelphia: Franklin Institute.

Larkin, J. (1980). Teaching problem solving in physics: The psychological laboratory and the practical classroom. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.

Larkin, J. & Rainard, B. (1984). A research methodology for studying how people think. *Journal of Research in Science Teaching, 21*(3), 235-254.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 69*, 344—354.

Newell, A. St Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Novak, J. D. (1984, April). *Learning science and the science of learning*. Paper presented at the American Educational Research Association annual convention. New Orleans.

Reif, F. (1980). Theoretical and educational concerns with problem solving: Bridging the gaps with human cognitive engineering. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.

Scandura, J. M. (1977). *Problem solving: A structural/process approach with instructional implications*. New York: Academic.

Schwab, J. J. (1984). *Science, curriculum, and liberal education*. Chicago: University of Chicago Press.

Shulman, L. S. & Carey, N. B. (1984). Psychology and the limitations of individual rationality: Implications for the study of reasoning and civility. *Review of Educational Research, 54*(4), 501-524.

Simon, H. A. (1978). Information processing theory of human problem solving. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes. Vol, 5. Human information processing*. Hillsdale, NJ: Erlbaum.

Simon, H. A., Langley, P. W., & Bradshaw, G. L. (1981). Scientific discovery as problem solving. *Synthese, 47*, 1-27.

Smith, C. 0. (1971). *The structure of intellect processes analvses system. A technique for the investigation and quantification of problem solving processes*. Unpublished doctoral dissertation. Houston, TX: University of Houston.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69*, 344-354.

APPENDIX A

TALK-ALOUD INSTRUCTIONS TO BE READ TO STUDENTS

Begin with informal chat with student.

Instructions:

I'm going to ask you to solve some problems about biology. The problems are presented as a part of a computer game. You can solve these problems in any wav you like. In fact, I'm not even interested in whether you get the right answers. But I am interested in how you think, the processes you use. Therefore, I'm going to ask you to talk or think aloud as much as you can. I'm going to turn on the tape recorder to make a record of what you are saying. [TURN ON RECORDER]

You won't have to do anything special. Just be yourself and work as you normally would. Some people say that they "mumble to themselves" when they are solving problems. If that's what you do, then all you have to do is mumble louder. In any case, try to talk constantly. Say what you are thinking and doing, even if it doesn't make sense.

First, I'll give you a problem just for practice. This problem is not like the problems you will have on the computer. Its purpose is to practice the talk-aloud technique. Do you have any questions?

Okay, work this one now and talk as much as possible.

APPENDIX B

PERMITTED COMMENTS DURING VERBAL PROTOCOLS

uhn-huh

okay

please, talk more it you can talk a bit louder, please can you say what you're thinking

mmmm

please tell me what you're writing remember, say what you're thinking

TECHNICAL COMMENTS

press <return>

press <P> or <G>

type in your id number tor this study

read the instructions

are you going to make a guess?

be sure to press <G> before you make a guess

press the <space bar>

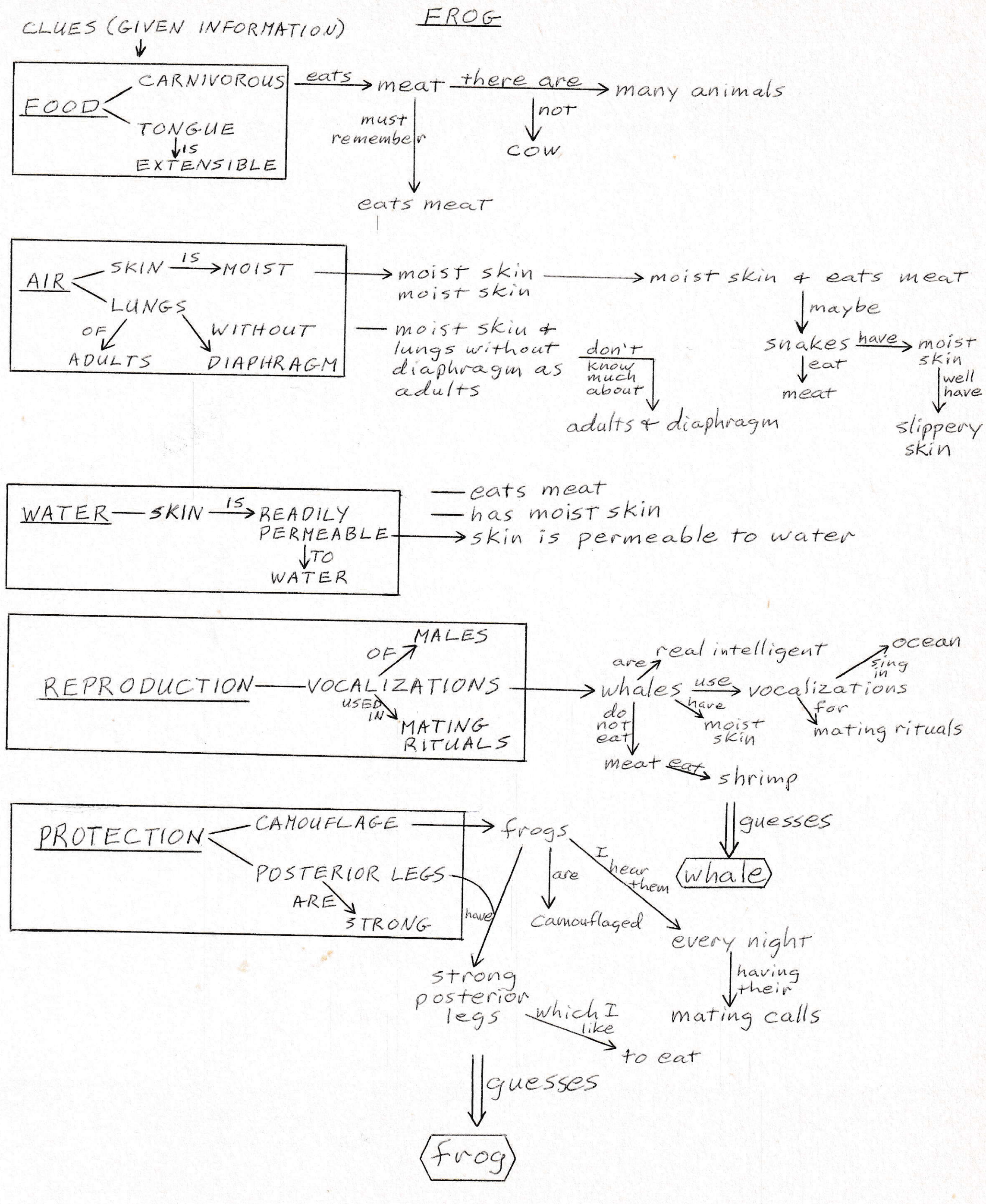

FROG
CLUES (GIVEN INFORMATION)
FOOD
CARNIVOROUS
TONGUE
IS
EXTENSIBLE
eats
meat
there are
many animals
not
cow
must remember
eats meat
AIR
SKIN
IS
MOIST
LUNGS
OF
ADULTS
WITHOUT
DIAPHRAGM
moist skin
moist skin
moist skin & eats meat
maybe
snakes
have
moist skin
eat
meat
well have
slippery skin
moist skin & lungs without diaphragm as adults
don't know much about
adults & diaphragm
WATER
SKIN
IS
READILY PERMEABLE
TO
WATER
eats meat
has moist skin
skin is permeable to water
REPRODUCTION
VOCALIZATIONS
OF
MALES
USED IN
MATING RITUALS
are
real intelligent
whales
use
vocalizations
sing in
ocean
for
mating rituals
have
moist skin
do not eat
meat
eat
shrimp
guesses
whale
PROTECTION
CAMOUFLAGE
POSTERIOR LEGS
ARE
STRONG
frogs
are
camouflaged
I hear them
every night
having their
mating calls
have
strong posterior legs
which I like
to eat
guesses
frog

FIGURE 1

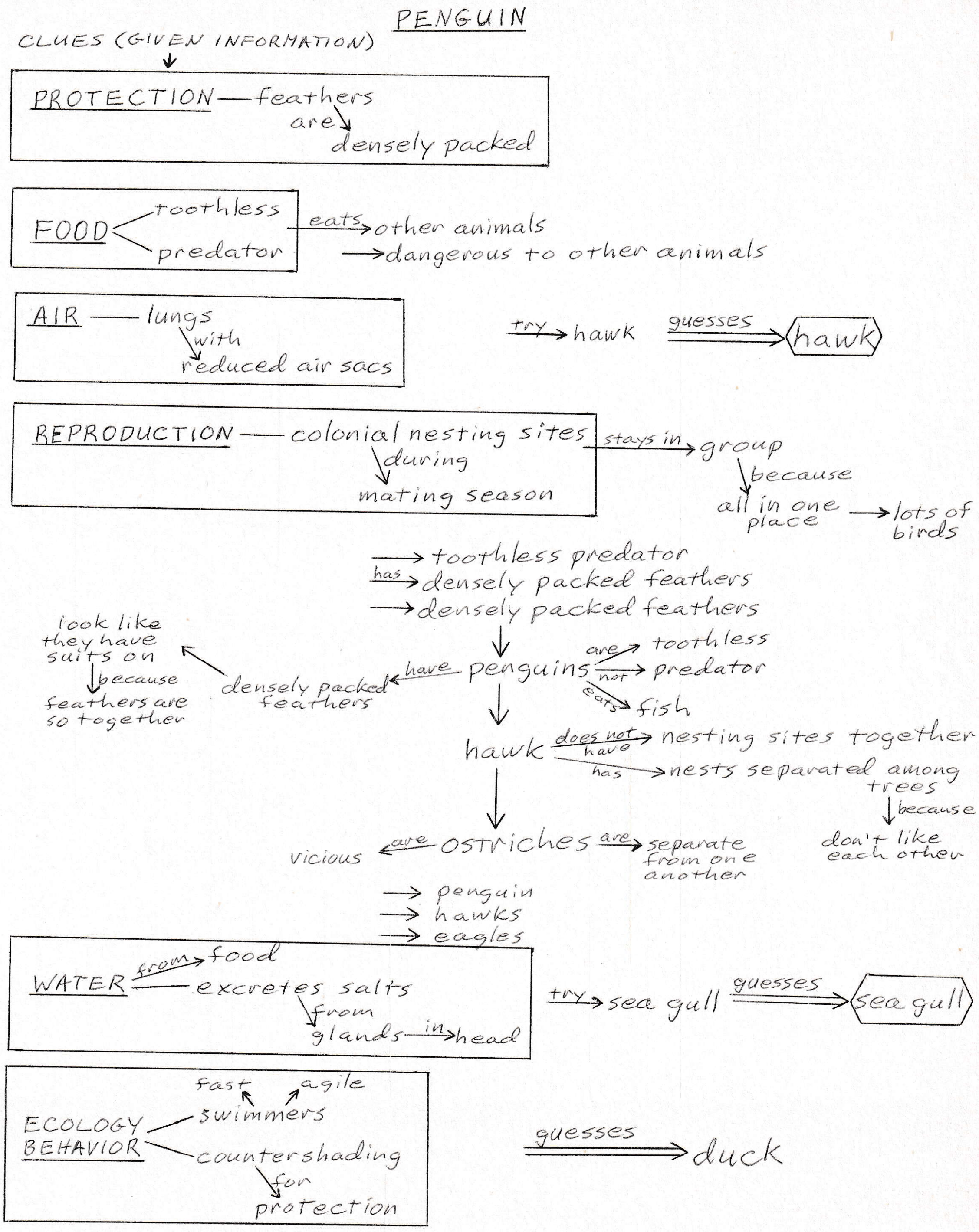
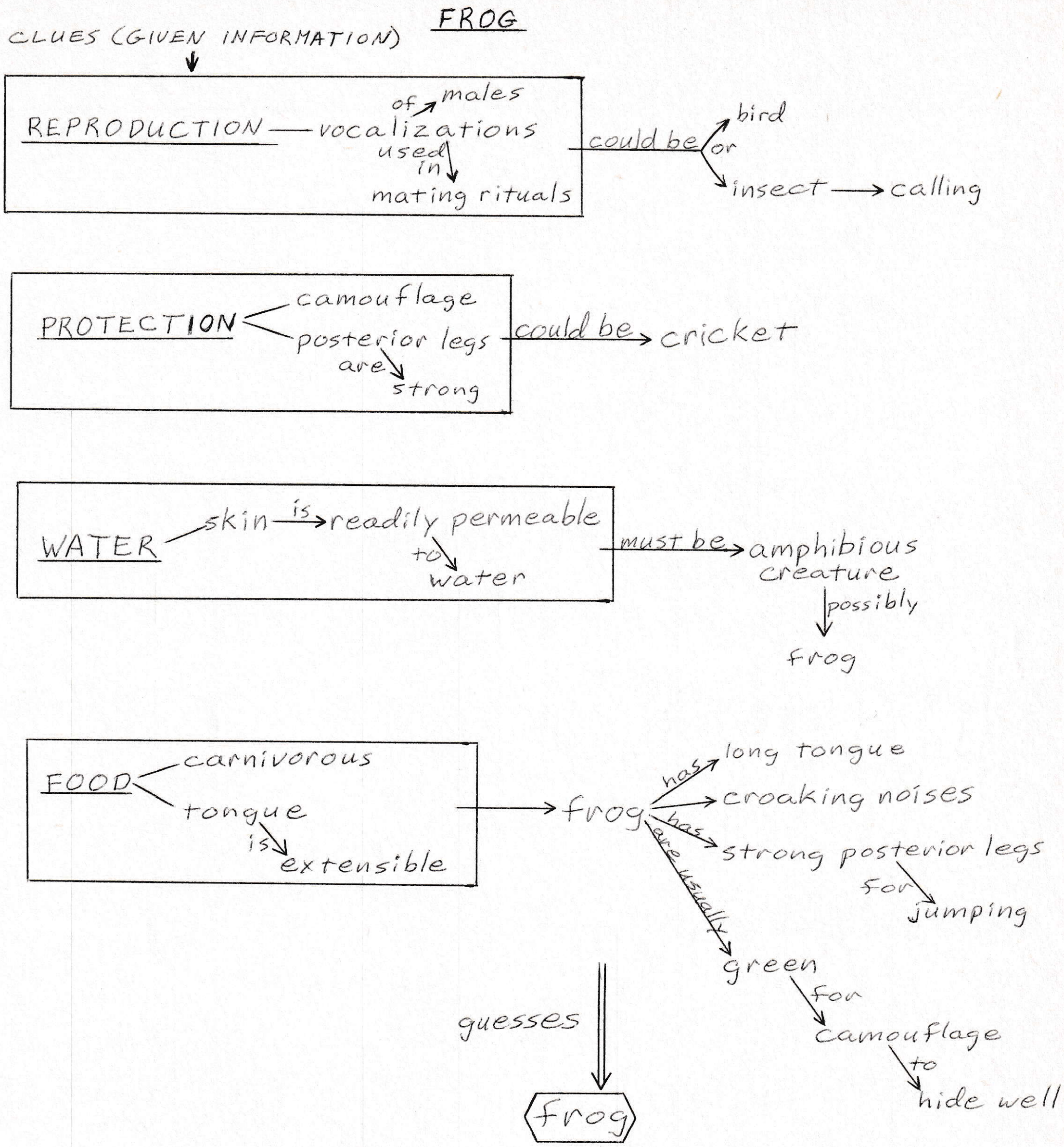

PENGUIN
CLUES (GIVEN INFORMATION)
PROTECTION — feathers are densely packed
FOOD — toothless, predator
eats other animals
dangerous to other animals
AIR — lungs with reduced air sacs
try hawk
guesses
hawk
REPRODUCTION — colonial nesting sites during mating season
stays in group
because
all in one place
lots of birds
toothless predator
has densely packed feathers
densely packed feathers
look like they have suits on
because
feathers are so together
have densely packed feathers
penguins
are toothless
not predator
eats fish
hawk
does not have nesting sites together
has nests separated among trees
because
don't like each other
vicious
are ostriches
are separate from one another
penguin
hawks
eagles
WATER — from food; excretes salts from glands in head
try sea gull
guesses
sea gull
ECOLOGY BEHAVIOR — swimmers: fast, agile; countershading for protection
guesses
duck

FIGURE 2



FROG
CLUES (GIVEN INFORMATION)
REPRODUCTION
vocalizations
of
males
used in
mating rituals
could be
bird
or
insect
calling
PROTECTION
camouflage
posterior legs
are
strong
could be
cricket
WATER
skin
is
readily permeable
to
water
must be
amphibious creature
possibly
frog
FOOD
carnivorous
tongue
is
extensible
frog
has
long tongue
croaking noises
has
strong posterior legs
for
jumping
are usually
green
for
camouflage
to
hide well
guesses
frog

FIGURE 3

PENGUIN

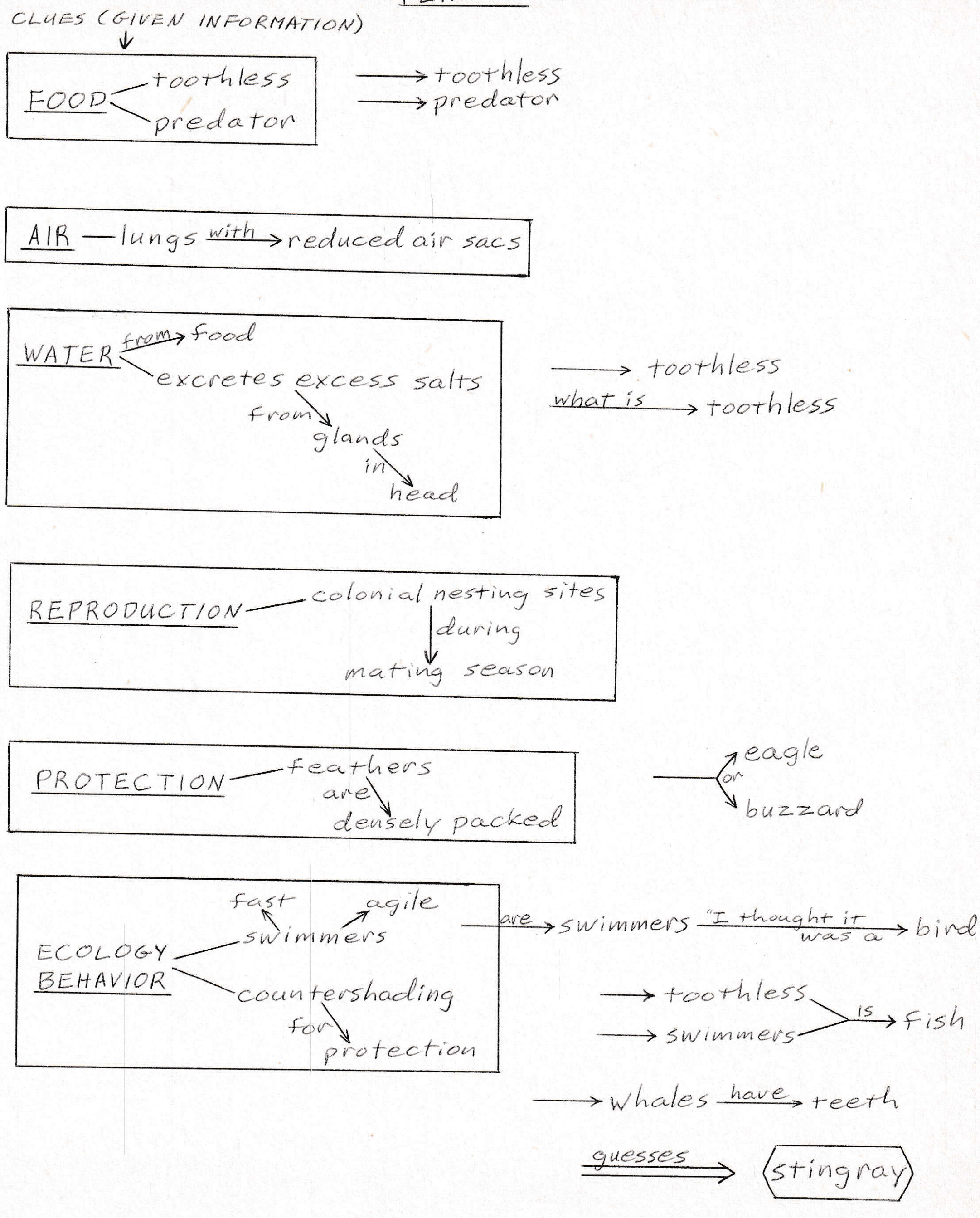


FIGURE 4

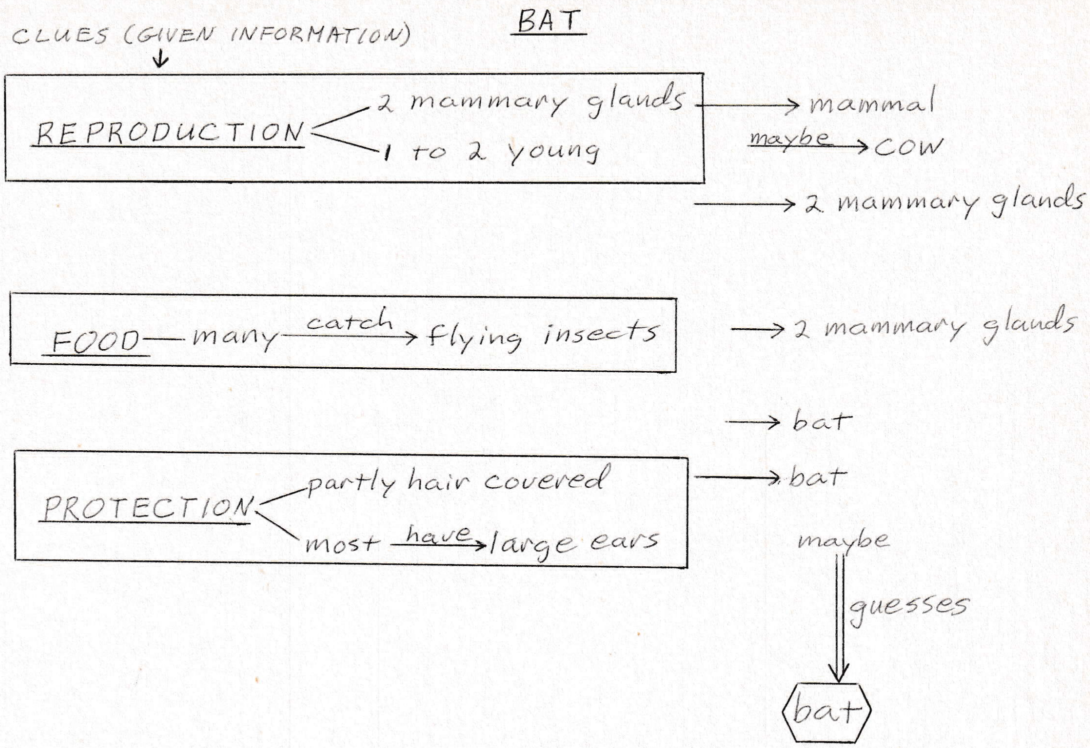

BAT
CLUES (GIVEN INFORMATION)
REPRODUCTION
2 mammary glands
1 to 2 young
mammal
maybe
COW
2 mammary glands
FOOD
many
catch
flying insects
2 mammary glands
bat
PROTECTION
partly hair covered
most
have
large ears
bat
maybe
guesses
bat

FIGURE 5

CORRELATION MATRIX OF COMBINED TOTAL SCORES

n = 10

	Mean	SD	Pos Cnj	Eli Cnj	Tot Cnj	Cor Rpt	Inc Rpt	Tot Rpt	Prx Inf	Dis Inf	Mis Inf	Tot Inf	Weak Test	Ver Cnj	Ref Cnj	Tot Test	Tot Dir Thk	Succ Dir Thk	Unsu Dir Thk	Degr Succ Scor	Gregorc Style Delineator			
																					CS	AS	AR	CR
Pos Con	.60	.41	***1.00	.19	***1.00	*.66	*.57	**.74	.29	-.38	-.31	.09	.46	.33	*.58	.44	**.82	.44	***.95	**.86	-.53	.12	.19	.26
Eli Con	.01	.02	.19	***1.00	.23	*.58	-.19	.50	.02	-.22	.10	.02	.19	.33	.33	.31	.37	.50	.18	-.02	.11	-.24	-.27	.26
Tot Con	.61	.41	***1.00	.23	***1.00	*.69	*.56	**.76	.29	-.39	-.30	.09	.46	.34	*.59	.46	**.84	.46	***.94	**.85	-.52	.11	.17	.27
Cor Rep	.51	.33	*.66	*.58	*.69	***1.00	.22	***.97	.48	-.28	-.10	.30	**.74	**.80	*.69	**.79	***.92	***.88	*.72	*.60	-.27	.21	-.17	.22
Inc Rep	.03	.08	*.57	-.19	*.56	.22	***1.00	.44	.14	-.01	-.13	.07	*.56	.34	**.77	.52	*.56	.11	**.83	.26	-.17	.54	.41	-.53
Tot Rep	.54	.35	**.74	.50	**.76	***.97	.44	***1.00	.47	-.26	-.12	.29	**.81	**.82	**.82	**.85	***.98	**.84	**.86	*.61	-.29	.32	-.06	.08
Pro Inf	.14	.12	.29	.02	.29	.48	14	.47	***1.00	.35	*.61	***.96	.28	.28	.18	.26	.52	**.77	.40	.24	.08	.22	-.35	.03
Dis Inf	.01	.02	-.38	-.22	-.39	-.28	-.01	-.26	.35	***1.00	*.57	*.56	-.21	-.14	-.14	-.17	-.22	-.01	-.21	-.39	*.60	.04	*-.60	-.12
Mis Inf	.03	.04	-.31	.10	-.30	-.10	-.13	-.12	*.61	*.57	***1.00	**.79	-.30	-.26	-.18	-.27	-.13	.26	-.11	-.45	.38	-.13	-.31	-.02
Tot Inf	.17	.15	.09	.02	.09	.30	.07	.29	***.96	*.56	**.79	***1.00	.10	.13	.08	.10	.33	*.64	.24	.01	.24	.14	-.43	.01
Wk Tes	.06	.09	.46	.19	.46	**.74	*.56	**.81	.28	-.21	-.30	.10	***1.00	***.94	**.84	***.98	**.80	*.64	**.82	.32	-.16	*.61	.18	-.42
Ver Con	.09	.16	.33	.33	.34	**.80	.34	**.82	.28	-.14	-.26	.13	***.94	***1.00	**.79	***.98	**.76	*.69	*.66	.29	-.08	*.57	-.08	-.26
Ref Con	.05	.07	*.58	.33	*.59	*.69	**.77	**.82	.18	-.14	-.18	.08	**.84	**.79	***1.00	***.89	**.82	.49	***.92	.33	-.18	*.57	.14	-.34
Tot Tes	.20	.31	.44	.31	.46	**.79	.52	**.85	.26	-.17	-.27	.10	***.98	***.98	***.89	***1.00	**.82	*.66	**.80	.33	-.13	*.60	.04	-.33
Tot DTh	1.51	1.00	**.82	.37	**.84	***.92	*.56	***.98	.52	-.22	-.13	.33	**.80	**.76	**.82	**.82	***1.00	**.80	***.93	*.67	-.31	.37	-.01	.03
Suc DTh	1.83	1.11	.44	.50	.46	***.88	.11	**.84	**.77	-.01	.26	*.64	*.64	*.69	.49	*.66	**.80	***1.00	.58	.34	.02	.18	-.29	.07
Uns DTh	1.34	1.07	***.95	.18	***.94	*.72	**.83	**.86	.40	-.21	-.11	.24	**.82	*.66	***.92	**.80	***.93	.58	***1.00	**.83	-.37	.48	.30	-.32
DgS Sco	.45	.25	**.86	-.02	**.85	*.60	.26	*.61	.24	-.39	-.45	.01	.32	.29	.33	.33	*.67	.34	**.83	***1.00	*-.64	.06	.01	.53
Con Seq	26.2	4.49	-.53	.11	-.52	-.27	-.17	-.29	.08	*.60	.38	.24	-.16	-.08	-.18	-.13	-.31	.02	-.37	*-.64	***1.00	-.29	-.53	-.30
Abs Seq	22.8	3.49	.12	-.24	.11	.21	.54	.32	.22	.04	-.13	.14	*.61	*.57	*.57	*.60	.37	.18	.48	.06	-.29	***1.00	.23	*-.61
Abs Ran	26.7	3.68	.19	-.27	.17	-.17	.41	-.06	-.35	*-.60	-.31	.43	.18	-.08	.14	.04	-.01	-.29	.30	.01	-.53	.23	***1.00	-.42
Con Ran	24.3	5.01	.26	.26	.27	.22	-.53	.08	.03	-.12	-.02	.01	-.42	-.26	-.34	-.33	.03	.07	-.32	.53	-.30	*-.61	-.42	***1.00

* P < .05   ** P < .01   *** P < .001

FIGURE 6

TOTALS OF COMBINED, SUCCESSFUL, AND UNSUCCESSFUL THINKING

CORRELATED WITH DEGREE OF SUCCESS SCORE

	Mean	SD	R	N
Degree of Success Score	.45	.25		10
Combined Total Directed Thinking	1.51	1.00	* .67	10
Successful Total Directed Thinking	1.83	1.11	.34	10
Unsuccessful Total Directed Thinking	1.34	1.07	** .83	9
Combined Total Conjectures	.61	.41	** .85	10
Successful Total Conjectures	1.00	.50	.18	10
Unsuccessful Total Conjectures	.42	.41	* .81	9
Combined Different Conjectures	.32	.21	* .79	10
Successful Different Conjectures	.48	.24	.09	10
Unsuccessful Different Conjectures	.27	.24	** .88	9
Combined Total Repeats	.54	.35	* .61	10
Successful Total Repeats	.40	.49	* .58	10
Unsuccessful Total Repeats	.54	.33	** .87	9
Combined Productive Thinking (Inferences)	.17	.15	.01	10
Successful Productive Thinking (Inferences)	.16	.24	-.07	10
Unsuccessful Productive Thinking (Inferences)	.16	.17	-.13	9
Combined Testing of Conjectures	.20	.31	.33	10
Successful Testing of Conjectures	.27	.42	.04	10
Unsuccessful Testing of Conjectures	.22	.33	** .86	9

R=.80 w/out Case #1 (beside Combined Testing of Conjectures)

* P < .05   ** P < .01   *** P < .001

FIGURE 7

GRAPH OF MEANS AND STANDARD DEVIATIONS OF TYPES OF THINKING

ON COMBINED, SUCCESSFUL, AND UNSUCCESSFUL TRIALS

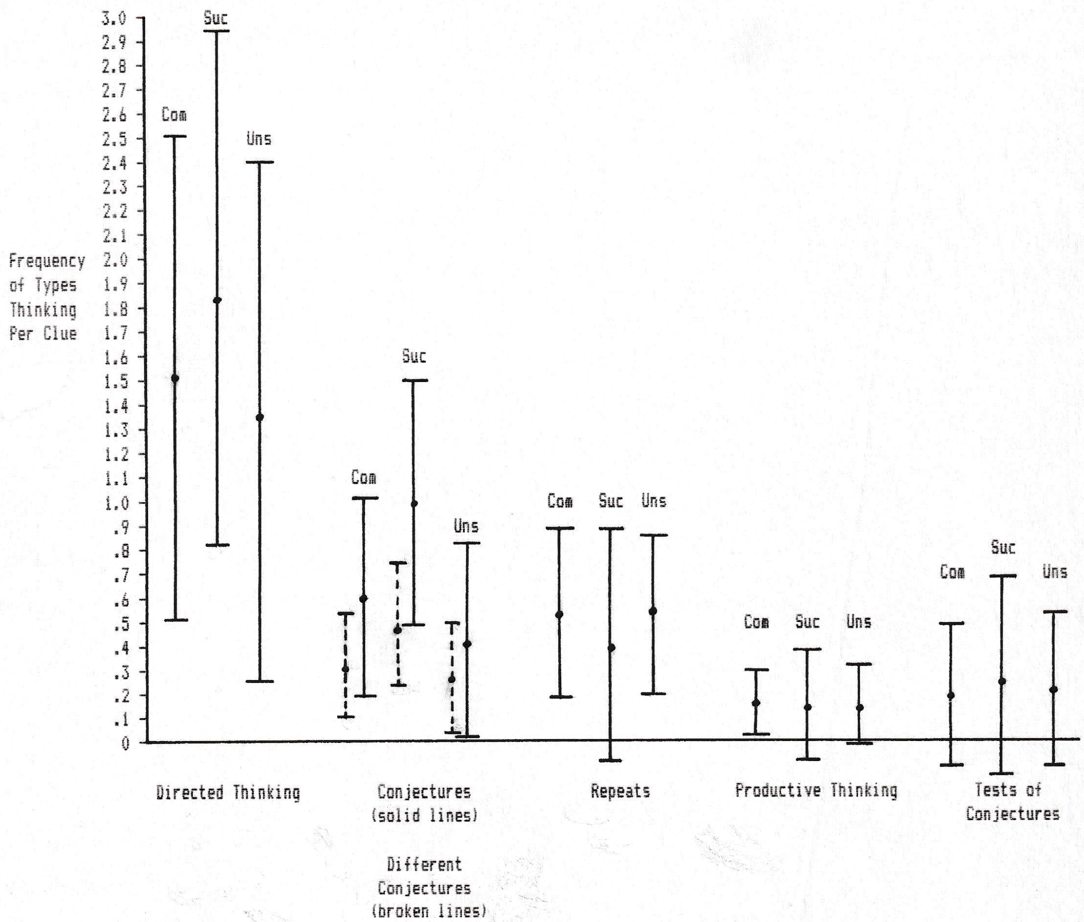


FIGURE 8