A majority of the empirical studies on questioning has focused on the effects of prior knowledge on the number and type of questions asked. Research has not been able to find support for the frequency hypothesis and the specificity hypothesis in questioning. The focus of these studies is briefly criticized, and a detailed critical assessment is made of their methods of study. It is argued that many studies on questioning have not succeeded in creating a realistic context of inquiry. Certain minimum conditions are outlined for setting up research on questioning. Also reviewed are the problems with: (1) using the same kinds of questions found in natural interactions; (2) creating a naturalistic exchange between questioner and respondent; and (3) enabling people to use other inquiry strategies besides questioning. It is argued that these issues merit attention because of methodological concerns and because they are likely to lead to important insights about questioning. A 38-item list of references is included. (SLD)
A critique of research on questioning

Hans van der Meij
University of Twente, Department of Education
P.O. Box 217, 7500 AE Enschede
Netherlands

Running Head: Questioning Research

Some of the ideas presented in this paper were published in the following article:
Abstract

A majority of the empirical studies on questioning has focused on the effects of prior knowledge on the number and type of questions asked. The research has been unable to find support for the frequency and the specificity hypothesis in questioning. This paper briefly criticizes the focus of these studies and gives a detailed critical assessment of their methods of study. It argues that many studies on questioning have not succeeded in creating a realistic context of inquiry. This paper outlines certain minimum conditions for setting up such research on questioning. A discussion is given of the problems with using the same kinds of questions as found in natural interactions, with creating a naturalistic exchange between questioner and respondent, and with enabling people to use other inquiry strategies besides questioning. It is argued that these issues merit attention not only because of methodological concerns but also because they are likely to lead to important insights about questioning.
It takes skill and motivation to ask a question. First, a person who generates a question must experience a perplexity and feel compelled to find things out. Second, expressing the question rightly is a difficult communicative task and must be handled carefully in order to avoid having a painful social encounter or facing public embarrassment. Third, the person must reflect about the success of the question asking strategy and the improvements that might be made in future. In short, what is sometimes considered a common part of a person’s behavioral repertoire is really a most remarkable achievement. Asking a question is an expression of will, of effort, of courage, of commitment, of skill, of knowledge and of lack of knowledge (compare Dillon, 1990).

Researchers have long been intrigued by what happens when someone asks a question. Generally they have found the matter a difficult one to investigate. Empirical studies have concentrated on the effects of prior knowledge on questioning (i.e. number and kinds of questions asked). More particularly, the research has tried to verify two hypotheses: a frequency hypothesis and a specificity hypothesis. Unfortunately, little support has been found. This paper suggests that this may be due to the fact that researchers have not paid enough attention to important problems in creating a realistic context of inquiry. A close look at these problems reveals several deficiencies, all of which are vital components for developing a theory of questioning (compare Dillon, 1986, 1990; Flammer, 1981, 1986; Usunov, 1987). The attention given to these problems therefore extends the pure methodological concerns. This paper first presents the two dominant research hypotheses and methods of study. Then, the various problems in creating a realistic context of inquiry will be discussed.

The frequency and specificity hypothesis

How is it that people are capable of asking questions about things that they do not know? This complex question has been at the basis of many studies on questioning. Unfortunately, researchers have often transformed this question into the much simpler notion of “How much prior knowledge is needed for a maximum number of questions to appear?” For example, Miyake & Norman (1979) have suggested that people ask the most questions when they face a task of moderate difficulty, given average prior knowledge. The task then fits their level of knowledge; the unknown is a knowable unknown. Too easy or too difficult tasks will lead to fewer questions. In the first condition there is no reason for asking questions and in the second there is no schema or framework for generating them.

Little support has been found for this frequency hypothesis. With the exception of their own study and part of Allwinn’s (1988), all studies have shown that people ask more questions for more difficult tasks (Flammer, Kaiser & Lüthi, 1981; Flammer, Kaiser & Müller-Bouquet, 1981; Flammer, Grob, Leuthardt & Lüthi, 1984; Fuhrer, 1986, 1987, 1989; Nelson-LeGall & Glor-Scheib, 1985; Smith, Tykodi & Mynatt, 1988; Van der Meij, 1990a, 1990b, 1990c).

As for predicting the kinds of questions asked, the prevalent view is that there will be fewer specific questions for more difficult tasks. According to this specificity hypothesis, a novice who tries to master a new computer system is likely to ask questions that are very broad (e.g. “What do I do?”), whereas experienced users ask more focused questions (e.g. “How do I merge this file?”).

There has almost been no support for the hypothesis (Allwood & Eliasson, 1988; Flammer et al., 1981, 1984; Fuhrer, 1986; Miyake & Norman, 1979; Van der Meij, 1990a, 1990b, 1990c). Only Smith, Tykodi & Mynatt (1988) found some support; however, the effect was limited to the first stage of a learning process. Novices who gained just a bit of competence quickly began asking the same kinds of questions as more experienced subjects.

Creating a realistic context

In spite of the fact that the frequency and the specificity hypothesis have not had much experimental support they have hardly been challenged. Miyake & Norman’s (1979) study, for example, is still the most widely cited research on the question asking frequency. This is probably so because people have found the idea intuitively appealing that knowing too little or knowing too much is detrimental to asking a question. Why then has the research consistently pointed out otherwise? And why doesn’t higher prior knowledge lead to more specific questions? In short, why has the research done so poorly in predicting question asking in the experimental setting?
Invariably, researchers have brought forward methodological arguments. For example, they have argued that prior knowledge was not measured adequately, although measures of general ability, of verbal intelligence, and of general and specific domain knowledge have been used. Also, occasionally it is reported that the experimental subjects varied too little in prior knowledge, or that the study has not yielded enough questions for detailed analyses (see Flammer et al. 1984; Fuhrer, 1986; Miyake & Norman, 1979; Van der Meij, 1990a, 1990b).

From the standpoint of cognitive psychology the focus on frequency and specificity is a rather poor one because the two variables are rather unspecific. There is much more to look at. For example, it would seem important that research on questioning addresses issues such as the quest for declarative (episodic or conceptual) knowledge versus procedural knowledge, that it studies how problems are transformed into questions and hypotheses, that it studies how questions are sequenced and that it investigates under what conditions people ask deep, complex questions rather than simple, obvious ones (see Scardamalia, & Bereiter, 1991).

In addition to shifting the focus towards deeper understanding of the cognitive processes that underly questioning (and changing the dependent variables) more attention should also be paid to the inquiry conditions in the experiments. One way or another, researchers try to create a situation that is in a sense a proxy to a real context of inquiry. They must therefore face at least three major problems: (1) The questions asked in the experiment should resemble the questions asked in reality. Unfortunately, it is quite difficult to create a setting for inquiry such that the questions that are asked reflect those that are found in a natural situation; (2) The exchange between questioner and respondent should be realistic. However, it is almost impossible to create both a natural kind of interaction and not to interfere with the knowledge that is tapped from questioners; (3) Question asking is only one way of gaining information among many. But, almost inevitably, experimental subjects are often confronted with a situation that limits their inquiry behavior to asking questions. Each of these three issues will be discussed below. The examples come mostly, but not exclusively, from research on question asking and prior knowledge.

Using the Same Kinds of Questions
It is critically important to study the same kinds of questions in the experiment as found in a natural situation. Experiments that use a question selection procedure (which is a method in which subjects select their questions from a prearranged set instead of generating them, see e.g. Nelson-LeGall, 1987; Nelson-LeGall & Glor-Scheib, 1985, 1986; Trope & Bassok, 1983; Trope, Bassok & Alon 1984; Van der Meij, 1990a, 1990b, 1990c) should therefore use questions found in the natural setting. The importance of this simple rule of hand can be illustrated with the research of Snyder (Snyder & Swann, 1978; Snyder, 1981). In their research Snyder and colleagues found that people (i.e. lay-interviewers) tend to search for information that confirms their beliefs. Trope & Bassok (1983) criticized this finding, arguing that the list of questions from which the subjects selected their questions did not reflect the kinds of questions people would normally generate for an interview. With a new list, based on questions asked in real interviews, they found question asking to be diagnostic rather than confirmatory (see also Skov & Sherman, 1986). There is also other evidence of the troublesome nature of the selection procedure. For example, when I compared question selection with question production in the same experiment he found little correspondence on matters such as frequency and type of question asked Van der Meij, 1990b). In contrast, when I correlated question-asking frequency (i.e. selection) in the experiment with question-asking frequency in the classroom he found a low but significant correlation (r = .28, p < .05).

The selection procedure is not a bad method by nature. It seems, for example, a practical way of gaining insight into the processes involved in question posing without having to bother with issues of question raising (i.e. generating). In addition, it bypasses the hard problem of the immediate categorization of the questions and of choosing the right answer during the experiment (Flammer, 1986). However, it does have at least one serious drawback and that is that it is very difficult to settle upon the right kinds of questions for the experiment. Unfortunately, researchers often cannot use the naturally occurring questions because there are frequently not enough of such questions, or because there is too little correspondence with the experimental task. What researchers have therefore done is that they generate experimental questions that resemble real questions. This resemblance is based upon two critical assumptions. First, each question is
categorized into a particular class of questions. Second, each class of questions is believed to correspond closely to a particular question function (see Nelson-LeGall, 1987; Nelson-LeGall & Glor-Scheib, 1985, 1986; Nelson-LeGall, & Jones 1990; Nelson-LeGall, Kratzer, Jones, & DeCooke, 1990; Van der Meij, 1990a, 1990b, 1990c). For example, Nelson-LeGall & Glor-Scheib (1985) classified some children's questions (e.g. "How do I solve 14 x 3?" and "Can you show me how to multiply fractions?") as requests for hints, and others (e.g. "What is 14 x 3?" and "Can you tell me the answer?") as requests for answers. Each question class was believed to reflect a certain information-seeking function. The hints were considered a good measure of question asking for the purpose of learning how to solve problems. The answers supposedly came from the child's intention to complete tasks without much personal involvement.

Both assumptions may be troublesome. First, categorizing questions into a particular class is a tricky business and researchers should probably go beyond getting good interrator agreement scores to assure the validity of the coding. Second, it may not be such a good idea to believe that there is a one-to-one correspondence between the form and function of a question. Perhaps assuming complex relationships is the better option. In any case, additional measures to the questions being asked should be used to substantiate these presupposed relationships.

For example, many researchers have tended to concentrate on the information-seeking function of questions. In dense social settings such as classrooms this may be the wrong focus. In these settings questioning might primarily serve the social purpose of helping pupils to develop relationships as the following example illustrates (Gumperz, 1981, pp. 18-19):

"The black children tended to call the teacher or the aide to ask for help and even after an explanation they were often heard to comment 'I can't do this' or 'I don't know how to do this'. All such sentences were pronounced with similar intonation contours, characterized by high pitch register, sustained tone, and vowel elongation on the last syllable. We (...) asked a group of black judges whether they thought the children really didn't know or didn't want to cooperate. The judges agreed in saying what the children really mean to say in these cases is 'Help me; I don't like to work alone'."

One may thus wonder whether the two types of questions distinguished by Nelson-LeGall really exhaust the ones that are generally asked in school. For example, Meece, Blumenfeld & Hoyle, 1988 argue for the presence of a third major function namely an ego, or social goal (see Nelson-LeGall, 1990). Categorizing the questions as serving a particular kind of information-seeking or social function is also difficult (see Usunov, 1987). Should one use syntactical, semantic or pragmatic considerations to categorize a question? And is it fair to assume a close correspondence between the expression, class and function of a question? Researchers must come up with an answer to these questions regardless if they use a selection procedure or whether the subjects are free to generate their own questions.

Shaping the exchange
Laboratory studies on questioning always struggle with the exchange between questioner and respondent. In these exchanges there are problems with timing, norm-setting, and with the content of the response.

Timing is crucial in answering questions and in general respondents should answer questions promptly and not a very long time afterwards. In experiments on the effects of prior knowledge on questioning this is problematic since direct responses may activate dormant prior knowledge. Moreover, the responses are likely to lead to a unique pattern of responses and follow-up questions for each subject and thereby seriously compromise comparability between subjects. Thus, if the experiment simulates a naturalistic 'question - response - processing of response' type of interaction it becomes virtually impossible to analyze the effects of prior knowledge on questioning. To solve this problem researchers have handled the timing problem in the following ways: they have not answered the questions at all, or they have asked their subjects to pose or write down all of their questions before answering each individual one (e.g. Allwood, & Eliasson, 1988; Flammer, Grob, Leuthardt & Lüthi, 1984; Miyake & Norman, 1979). These solutions are unsatisfactory for a number of reasons. First, in the absence of direct responses questioners are unlikely to be concerned with communicating their precise informational needs. They need not create a 'common ground'. The effect may be an abundance of questions that are difficult to interpret (compare Miyake & Norman, 1975). Second, the subjects probably become more
selective in what they ask and they are more likely to keep their questions for themselves. When this leads to an increase of 'internal' questioning the validity of the study is seriously affected.

Third, since the subjects cannot base their questions upon responses to previous questions they are severely limited in following a plan of action; they simply cannot follow a strategy in which questions and responses are integrated. Instead, they must revert to a strategy of asking as many questions as possible, finding out only later that a number of these questions were useless in view of the answers received to critical previous questions.

All experiments follow the Gricean principle of cooperation. That is, the norms are set in favor of question asking and social constraints are minimized to give subjects maximal freedom to express questions as they arise. In addition to removing obstructing conditions, the instructions often explicitly encourage subjects to ask questions. In short, question asking is not merely enabled, it is clearly stimulated by the instructions (e.g. Allwood, & Eliasson, 1988; Fishbein, Eckart, Lauver, Van Leeuwen, & Langmeyer, 1990; Flammer, Kaiser & Lüthi, 1981; Flammer, Kaiser & Müller-Bouquet, 1981). How will this effect question asking? I would speculate that these procedures increase the chances that some subjects make up questions to please the experimenter. Other subjects may ask only a few questions because they feel a little uneasy about taking the floor while a 'superior' - the experimenter - is present (see Dillon, 1988, 1990; Mishler, 1975, 1978) or because they do not believe that the conditions for question asking are biased in their favor - there are ambiguous signs since timing is not arranged at the convenience of the subjects.

Another problematic part of the exchange lies in answering. How should one determine the content of the response? In all interactions there are difficulties with answering. For example, what should one do with questions that are based on a false presupposition as in the famous question "Have you stopped beating your wife?" And how should one respond to open questions like "What do I have to do?" In all studies the researchers must have formulated some rules of conduct. However, only Flammer and co-workers have described these rules in detail (see Flammer, Kaiser, & Lüthi, 1981). According to Flammer, respondents require knowledge about the following factors. First, they need domain or task-specific knowledge. Second, they must know how to classify the questions. Third, they must have a set of principles for selecting the right response. In his experiments, especially in the later ones, Flammer worked hard to standardize these factors. So he began with drawing a very detailed chart of the experimental task (i.e. preparing a mousse-au-chocolat) and presented this chart to the respondents who then all had the same task-specific knowledge. Next he categorized the domain into a number of classes (e.g. goals, actions, criteria for success, ingredients and instruments). For each domain-class he further specified the kinds of questions into types (e.g. yes/no, global/specific, one or more steps) and coupled each to a specific answer rule. So, if an experimental subject asked a question such as "With what thing do I do this?" all respondents would classify it as an instrument question of a global nature. Following the prescribed response rule they would answer "With a rod". Unfortunately this important information is left out of the published article of the experiment (Flammer et al. 1984). Interestingly, Flammer later (1986) argued that he found himself tangled in issues concerning question answering rather than in question asking in which he was interested. He has the good company of researchers in artificial intelligence research (e.g. Lehner, 1978), in philosophical research (e.g. Harrah, 1984), and in communication research (see Dillon, 1990) who all address answerhood. In contrast, most of the empirical studies on questioning do not describe the behavior of the respondents, nor the response rules upon which this is based (e.g. Fuhrer, 1989; Fishbein et al., 1990; Smith, Tykodi, & Mynatt, 1988). This is a serious omission. Without this information the findings are virtually worthless because it is hard to imagine what interactions have taken place.

Enabling other Information Seeking Strategies
A third problem in experimental research on questioning is that one would like to have subjects use their full range of information-seeking strategies in dealing with the problem at hand and not just question posing. Normally there are lots of ways of gaining information besides question posing. For example, people can consult study-guides, dictionaries, encyclopedia and other written materials, or they can touch or observe certain phenomena. Moreover, they can reflect upon a risen issue for a long time without ever expressing a question. Asking a question to a respondent
is thus but one of many ways in which people search for information. In most experiments, however, the subjects can only ask questions to find things out besides guessing and inferencing. The researchers have had good reasons for minimizing other inquiry behaviors. As the obstacles to asking questions are quite strong, researchers sometimes cannot do otherwise but to stimulate them by invitation. Also, it is often tacitly assumed that the expressed questions in the experiments reflect (a wide range of) the questions that underly other inquiry behaviors. Thus, if someone, say, consults a document it is presumed that the person has a question in mind for which he or she seeks an answer in the document. However, some questions just are too idiosyncratic, or ill-formed that respondents cannot be expected to give helpful answers. In such a case people are likely to use other inquiry behaviors than expressing a question. More generally then, since the subjects in most of the experiments can only ask questions, these studies reveal nothing of the relationships of question asking with other kinds of inquiry behavior.

Conclusion

In general, research on questioning and prior knowledge has not been very successful in verifying either the frequency hypothesis, or the specificity hypothesis. Besides little success, the cognitive research on questioning also has not paid enough attention to issues such as how problems are transformed into questions and how people become committed to a quest for knowledge.

Most of the published studies on questioning also do not pay enough attention to the problems associated with creating a realistic context of inquiry. More attention is needed on using the right kinds of questions, with following an exchange that allows for prompt and standardized responses, and with presenting alternative options for information-gathering strategies. Some of the cited studies have addressed a number of these problems, but unfortunately, their systematic approach can often be detected only by looking at a series of studies covering a long period of research. This is, for example, the case with the studies of Nelson-LeGall which cover approximately six years of research. Occasionally too, one finds such information in unpublished, internal reports that are difficult to obtain (i.e. Flammer's study).

The problems mentioned in this paper are not of interest solely because of methodological concerns. The approaches that researchers take will surely also contribute towards a deeper understanding of questioning and advance the development of a theory of questioning. Their stance with regard to issues such as the distinction between the form and function of questions and the distinction between information-seeking and social functions of questions may inform us about the problematical topic of defining questions. Standardization efforts with regard to the rules of conduct for the respondent can yield important insights into the pragmatic properties of question-answer situations. Also, knowing someone's reasons for posing a question as opposed to using another means of inquiry would seem to be an important building block in a theory of questioning.

References

   Verbale Informationssuche. Frankfurt am Main: Lang.
   Questioning. In O. Hargie (Ed.), A handbook of communication skills (pp. 95-127). London: Croom Helm.
   Learner's questions and comprehension in a tutorial setting. Journal of Educational Psychology, 82, 163-170.


Asking how to act. *Archives de Psychologie*, 52, 103-120.


Effects of social density and pre-knowledge on question asking in a novel setting. *Journal of Environmental Psychology*, 7, 159-168.

Effects of prior knowledge, crowding, and congruence of subjects’ and others’ goals on question asking in an unfamiliar setting. *Psychological Reports*, 64, 131-145.


The logic of questions. In D. Gabbay and F. Guentner (Eds.), *Handbook of philosophical logic* (pp. 715-764).


Student’s goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology*, 80, 514-523.


To ask a question one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, 18, 357-364.


Van der Meij, H. (1990b). Question asking: to know that you do not know is not enough. *Journal of Educational Psychology, 5*, 505-512.