Judging How Heavily a Question is Loaded: A Pragmatic Method

DOUGLAS WALTON

The purpose of this paper is to improve the curriculum in the teaching of informal logic and critical thinking by throwing new light on a problem that is quite important in philosophy generally, and in the fields of statistics, political science and law, as well-the problem of how to evaluate loaded questions. The problem of dealing with fallacious or otherwise tricky questions has, since ancient times, been a subject of some interest to philosophers, and it is a complex problem, with many different factors that need to be taken into account. The subject is also of current interest in relation to recent concerns about the widespread abuse of statistics in polls that deceptively look precise because they are scientific surveys. But the results are skewed one way or another by effects of the wording of the question (effects not measured by the announced estimates of the probability of error). The problem here is that the connotations of the words used in the question can produce a large variance in the statistical outcome of a survey. In this paper, the factor studied is that of how to judge whether and how a question is "loaded," in a way that it is important for a critical thinker to be aware of, and to deal with by crafting a careful reply, instead of giving a direct answer, or taking a given answer at face value.

In this paper, the method of judging how heavily a question is loaded is a method that is pragmatic in nature. The judgment is made by applying a normative model of dialogue to a particular case in which a question was asked. The normative model is prompted by the conversational context of the case, insofar as it can be determined from the question itself and the text of discourse surrounding it (as known in the case). The method is both normative and pragmatic in that it works by applying an abstract model of dialogue to the known particulars of question use in a given case. However, as will be shown, empirical tests (of the kinds used in statistical surveys) can be brought to bear as well.

By studying some classic cases of the use of loaded questions of the kind typically cited in textbooks on logic and critical thinking, it will be shown how the problem with such questions when they are fallacious (and otherwise significantly harmful from a viewpoint of critical thinking) is one of conceal-

© Inquiry: Critical Thinking Across the Disciplines, Winter, 1997. Vol. XVII, No. 2.

ment. The question, on the surface, appears to have a purely information-seeking function in the collection of data. But underlying the appearance, the question has a function of persuasion. The question is really an argument, or is being used as an argument for the purpose of persuasion. There is nothing intrinsically wrong with persuasion dialogue, but its concealment in questions can be a serious obstacle to the coming of epistemologically significant conclusions in the collection of data by polls and other epistemic methods based on the asking of questions.

1. Introduction

The traditional fallacy of many questions, sometimes also called the fallacy of complex question, has been treated in logic textbooks for many years using an old example that has an even more powerful impact in the current context of concern about domestic abuse.

CASE 1: HAVE YOU STOPPED ABUSING YOUR SPOUSE

It is not too hard to see why this question has been classified as a fallacy or sophistical tactic. As Hamblin (1970, p. 38) put it, the spouse abuse question "seems designed to force" innocent non-spouse-abusers "into admissions of guilt." Generally, it is not too difficult to get a grasp of how the tactic works. The syntactic structure of the question is that of a yes-no question (Harrah, 1984, p. 716), a type of question that admits of only two direct answers: 'yes' or 'no.' But no matter which of these answers is given by the respondent, he concedes having committed the crime of spousal abuse.

The study of the sophistical uses of complex, tricky questions has a long history. Eubulides, who lived around the time of Plato, was the inventor of many paradoxes (Kneale and Kneale, 1962, p. 114), including the paradox of the horned man, which can be conveyed by the following question.

CASE 2: HAVE YOU LOST YOUR HORNS?

The sophism in this question can be explained by observing that no matter which of the two direct answers is given by the respondent, he concedes that he had horns at one time, presumably a damaging admission (or one that makes him look ridiculous, anyway). How the trap works was even better explained using another example. Alexinus of Elis, a member of the school of Eubulides (Diogenes Laertius, *Lives,* II. 109), is said to have asked Menedemus, another philosopher, if he had left off beating his father. His answer was, "I was not beating him and have not left off." Alexinus is said to have insisted that the question should have been answered by a plain "yes" or "no" (*Lives,* 11. 135).

Studying such tricky verbal tactics of questioning may be taken to be an abstruse philosophical exercise in logic that has no application to the real world. But far from it; studies of so-called "response effects" by empirical researchers like Schuman and Presser (1981) have shown abundantly that

many questions used in statistical polls and surveys, where the outcomes are very significant in influencing public opinion on all kinds of matters important in public affairs, restrict a respondent's capability to give an answer that expresses his real opinion on an issue, in much the same pattern of biasing that the spouse abuse question indicates. Such questions can be complex, can contain presuppositions that the respondent doesn't agree with or even understand, and may leave out a "none of the above" option that is needed. Consequently, the question, as used in a poll, produces a response effect, a skewing of the response away from the response that represents the respondent's real opinion on an issue. As a result, crime statistics, unemployment statistics, and all the other kinds of statistical findings that are treated as being so important to decide public and personal deliberations, can be highly misleading. Misuse of such question in polls and surveys has in fact now become so widespread that Crossen (1994) has been able to cite data indicating that misuse of polling, sometimes called "push-polling," has become a standard method of influencing public opinion used by corporations, politicians, advertisers, and public relations firms. It has become a large and profitable business.

The method used in this paper to evaluate questions is pragmatic in nature, as mentioned above, and can generally be categorized as *dialectical,* meaning that a question is viewed as a contribution to a conversation or dialogue exchange between two parties, called the proponent (or questioner) and the respondent. Such a dialogue is goal-directed, the participants take turn asking and replying to questions, and the structure of the dialogue is normative, in the sense that it has procedural requirements, and only certain types of moves fit into the procedures that are appropriate, or useful to contribute to the goals. How these structures look, and what kinds of rules they have, can be appreciated by looking at Hamblin (1971), Mackenzie (1981), or Walton and Krabbe (1995). The general framework can be called problematological in the sense of Meyer (1995). According to the problematological conception of a question, the asking of the question presupposes an issue, a framework of dialogue in which doubts are being raised, or a contrary opinion is being expressed. Hintikka (1995) has used a formal method of evaluating why-questions that can be described as dialectical and problematological in nature. Hintikka formulates a pragmatic context for questions in which a given question plays a role in what is called an interrogative inquiry, a type of dialogue, as it would be called here. A "big question," or issue, is called the *principal question* of the inquiry. The goal of the inquiry is to answer this big question. Evaluating a particular question in a given case, called a "small question" by Hintikka (p. 643), is carried out examining the place the small question has in the larger framework of answering the big question.

Readers in speech communication will be familiar with the pragmatic and dialectical system of van Eemeren and Grootendorst (1992), in which questions are evaluated as contributions to a type of` dialogue called the critical discussion. The goal of a critical discussion is for the two participants to

resolve a conflict of opinions by means of a verbal, rational exchange of viewpoints. The critical discussion corresponds approximately to the persuasion type of dialogue described in Walton and Krabbe (1995). In a persuasion dialogue, there are two participants, a proponent and a respondent. Each has a thesis,to be proved, and what each must do to prove his or her thesis is to use valid (or at least structurally correct) arguments that have as premises propositions that have been accepted by the other side. The initial situation is a conflict of opinions, and the goal is to resolve (or at least throw light on) the conflict by articulating the strongest arguments on both sides, and the strongest criticisms of those arguments. The argumentation stage of the persuasion dialogue takes the form of questions posed by one party and replies given the other party. The parties take turns asking questions and putting arguments forward. Each party starts with some initial concession by the other side, and then, in a step by step sequence of argumentation, tries to use these concessions, and gain other concessions as well, in a chain of argumentation that will prove his or her thesis.

In a dialectical evaluation of the use of any argument (or other speech act, like a question) in a given case, how the argument is to be evaluated depends on the type of dialogue it was supposed to be a part of. In a persuasion dialogue, it is normal for a question to contain presuppositions that support the viewpoint of the questioner, or to even contain loaded terms that have strong persuasive connotations. But if the same question is being used in the scientific collection of data in a poll, it is supposedly part of an epistemic framework in which information is being collected, and then used to prove some conclusion, based on scientific evidence. The conclusion then is supposed to be established as knowledge, or at least as a proposition that is supported by scientific evidence based on empirical data. If such a question contains persuasive elements, like loaded terms that have a persuasive function, and no notice is taken of this persuasion effect inherent in the question, epistemological norms of scientific data collection will be violated. The question will be biased. Such a bias may not be a problem is a persuasion dialogue, where strong advocacy of one's viewpoint is normal, and even necessary for a good dialogue. But if the same question is being used to collect data in what is supposedly a scientific information-seeking epistemic framework, bias could be a significant problem, especially if it is concealed. The detection of such bias is an important goal of critical thinking.

The pragmatic method evaluates the asking of a question in a particular case in relation to the norms appropriate for the type of dialogue that the question was supposedly a part of. What is important are expectations about the conventional type of dialogue the questioner and respondent are supposedly engaged in. Apart from this base line, however, other information is required as well. We may also need to know what prior questions were asked, and how these were answered in the sequence of dialogue in the case. One particular tool is very valuable for collecting and analyzing such evidence. For

use in the analysis and evaluation of the kinds of problematic questions cited above as the object of this study, the analytical tool of choice - developed specifically in (Walton, 1989, pp. 67-70) to deal with the spouse abuse question - is the profile of dialogue. A *profile of dialogue* is an ordered sequence of moves (generally questions and replies) in a dialogue exchange between two parties, with an initial part, just prior to the designated move, and a subsequent part, just following the designated move. Presenting a profile of dialogue is a way of providing a normative model of the local context of a move - for the purpose of the present concern, the move would be a question - by placing it in relation to the immediately adjacent moves in a context of dialogue. As Krabbe (1992) has shown, the profile of dialogue is a useful tool for analysis and evaluation of a move made in a dialogue, where giving a formal reconstruction of the whole abstract normative structure of the dialogue as a whole (for example, as in the kinds of dialogue structures formalized in Walton and Krabbe, 1995) would be too tedious and time-consuming. So the profile is a kind of short cut way of representing the localized dialectical structure of a move made in argumentation.

The profiles method, as applied in Walton (1989, chapter 2), shows how the task of evaluating the spouse abuse type of question requires a pragmatic approach that is sensitive to the different contexts in which this kind of question can be used. For example, suppose that the respondent, in the given case, is the defendant in a trial, and he is being cross-examined in court. In the just prior sequence of question and replies in the cross-examination, he has admitted to the attorney that he has a spouse, and that he has abused that spouse in the past. In that context, asking the spouse abuse question is perfectly reasonable. Certainly, at any rate, the question, as used in this type of case, should not be classified as an instance of the fallacy of many questions. But by contrast, consider another kind of case where the respondent is not guilty of spouse abuse, and has not admitted to the questioner, in the prior dialogue, that he has abused his spouse. Here, asking the spouse beating question is problematic, in exactly the way that Hamblin describes (just above), and it is to describe its use in this very kind of case that the logic textbooks present it as a fallacy. It could be the very same question asked in both cases. What makes the one case fallacious and the other not is the difference in the context of use.

According to the analysis of the spouse abuse question given in Walton, (1995, pp. 202-205), the reason why such a question is rightly judged to be fallacious (when it is fallacious) as used in a given case, is that the question, as asked, balls up a sequence of questions. The profile of dialogue is used to show, on a case-by-case basis, how the fallacious question asks a series of questions in an incorrect order-incorrect in relation to the constructive sequence of questioning and replying that would contribute to the goal of the dialogue. The mismatch between the actual sequence and the requirements made clear by the profile is revealed by applying the profile to the actual

sequence of questioning and replying in the given case. The profile represents the right sequence of questioning and replying, and the deviations from the profile that occur in the actual case can be analyzed and evaluated in relation to the requirements of the normative model of dialogue.

2. Loaded Questions

For a full pragmatic analysis and evaluation of the kinds of problematic questions cited in section one above, the reader is referred to Walton (1989). But there is one aspect of this analysis that could be developed further. The questions treated in section one suffer from many problems that make them of interest to logicians. One problem is that the questions are complex, in the sense that they pose more than one question in the one question. Another, as explained above, is that they tend to force the respondent to give an answer that is somehow unsatisfactory as a reply. But the particular aspect of interest here is that these questions tend to be loaded questions, in the sense that they contain a bias or spin towards one side of some controversial issue. Clark and Schober (1992, p. 29) characterize loaded terms in questions, like "freedom fighters" versus "anti-government guerrillas" as loaded terms that set the perspective from which a question is to be answered. For Clark and Schober, loaded questions are an especially acute problem in statistical surveys, because they create response effects that bias the outcome of the survey in a way that is not measured by the numerical margins of error that are published with the results of the survey. So the loaded aspect of questions is important, but can be easily overlooked.

In Walton (1989, p. 18) a question is said to be loaded where the respondent is not committed to a presupposition of the question. A question is said to be aggressively loaded (argumentatively stacked) where it has a presupposition contrary to, or damaging to, the respondent's side of the argument. These definitions depend on the concept of commitment in dialogue-the pragmatic concept analyzed in Walton and Krabbe, (1995), and the concept of a presupposition of a question - which has been developed further in Walton (1995). So some updating of the concept of a loaded question is now in order.

The new approach to developing an operational account of the concept of a loaded question is more general, in that presents an integrated theory of loaded questions, loaded terms and loaded statements. Thus it draws together material on loaded questions with other material that has traditionally been treated by the logic textbooks under the headings of other kinds of fallacies. We begin with loaded statements.

A statement (proposition) is loaded if, and only if, it is used by a proponent in a dispute between two parties in such a way that it goes against the side of the respondent. By a dispute is meant an argumentative exchange between two parties in which the one party has the goal of proving a particular proposition

called her thesis, and the other party has the goal of proving the opposite proposition (in other words, the negation of the first party's thesis). The following case is an example of a loaded proposition used in a dispute.

CASE 3 : HELEN AND BOB ARE ARGUING ABOUT THE ABORTION ISSUE. HE ARGUES FOR THE PRO-LIFE SIDE AND SHE ARGUES FOR THE PRO-CHOICE SIDE. AT ONE POINT IN THE ARGUMENT, BOB ASSERTS THE FOLLOWING PROPOSITION : "ABORTION, THE KILLING OF A BABY, IS AN ACT OF SERIOUS MORAL CONSEQUENCE." HELEN REPLIES, "THE FETUS IS NOT A BABY."

In this case, the proposition "Abortion, the killing of a baby, is an act of serious moral consequence." is a loaded statement. It is loaded because it goes against Helen's side of the argument, in just the following way. The statement has strongly persuasive implications.

The proposition that Bob asserted in case 3 implies the proposition that the act of abortion is an act of killing a baby. As we know, this implied proposition is one that pro-choice advocates would strongly deny, and it is not hard to appreciate why. If it could be proved by arguments, then those arguments would already go quite a ways along the path required to prove that abortion is morally wrong. So in case 3, Helen would quite rightly (from a viewpoint of critical thinking) see the proposition Bob asserted as a loaded statement. It does not follow that Bob's statement is inherently inappropriate or logically inadmissible in the dialogue, for presumably, the dialogue that Helen and Bob are engaged in a persuasion dialogue. Still, Helen should have a right not to have to accept the statement, on the ground that it is loaded, and she has a right to point out that the statement is loaded, and to insist that it has a burden of proof attached to it.

Statements can be more heavily loaded or less heavily loaded. A statement is *heavily loaded* if and only if it goes quite a distance along the path towards proving the respondent's thesis false in a dispute. An example of a heavily loaded statement would be a case comparable to case 3, except that Bob had said, "Abortion, the act of murdering a baby, is an act of serious moral consequence," in place of the assertion he made in case 3. This statement can be said to be heavily loaded, in the context of the dispute between Helen and Bob, because if true, it would directly imply Bob's thesis in the dispute - the proposition that abortion is an immoral act (an act that should not be carried out).

A term that occurs in a proposition is a *loaded term* if, and only if, the proposition is loaded, and would cease to be loaded (or so heavily loaded) if that term were removed from the proposition and replaced by a comparatively neutral, or less heavily loaded term. For example, the term 'murder' in the example just above, can be said to be a loaded term, because if it were replaced by the less heavily loaded term 'killing', the proposition Bob asserted would be less heavily loaded. Also, the term 'killing', as used in case 3, can be said to be a loaded term, because if it were replaced by the neutral, or less heavily loaded

term 'terminated,' the proposition Bob asserted would be less heavily loaded. Of course, the term 'baby' is also a loaded term, as used in the context of the abortion dispute in case 3, because replacement of it with a term like 'fetus' would make the resulting proposition less heavily loaded.

Finally we get to loaded questions. A question is said to be loaded, in the sense best suited to the purpose of this investigation, if it contains terms or propositions that are loaded. What exactly'contains' means here is a subject of some controversy, however. Traditionally what is taken to be referred to by this expression is the concept of presupposition, a concept that many different theories have been put forward in linguistics and philosophy to explain. The general idea is that a question, like the spouse abuse question, contains propositions that are presupposed by the question. But what is really needed for our purposes here is a pragmatic concept of presupposition that should be defined as follows. A proposition is *a presuppostion of a question* if and only if the respondent becomes committed to that proposition as soon as he gives any direct answer to the question. So, for example, with respect to the spouse abuse question, the two propositions, 'I (the respondent) have a spouse,' and'I (the respondent) have abused that spouse, are presuppositions of the question- meaning that the respondent immediately becomes committed to these two propositions as soon as he answers 'yes' or 'no' to the question. The concept of presupposition is defined here with reference to the commitments of the respondent in a dialogue exchange with a proponent (questioner).

3. How Heavily is the Question Loaded?

From a point of view of critical thinking, the most important thing about loaded questions, is for the respondent to recognize that the question is loaded, and to be aware that he may be entrapped, or at least concede damaging commitments that can later be used against him in a dialogue, if he gives an answer to the question. What the critical thinker needs to do, in such a case, is to question the question, and perhaps even to challenge it, by rebutting the propositions imputed by it. The respondent also needs to be careful, in some cases, not to attack the question too aggressively, for that response may also be taken to imply guilt for the accusation made in the question. These critical thinking skills can be taught without using any technical tests to determine how heavily loaded a question is, in a given case.

But still, it is interesting to explore the question of how to judge how heavily, a question is loaded. There are two methods - an empirical method and a normative method. The normative method is the more fundamental, but the empirical method is applicable because the questions that are the object of study are expressed in natural language. Terms used in such questions are loaded because they are words or phrases that have a conventional (lexical) meaning containing emotional connotations of the kind classified by Bentham (1838, 1962) as "laudatory" (positive) or "vituperative" (negative).

The normative method is based on the classification of the types of dialogue set out in Walton and Krabbe (1995, p. 66) and in Walton (1995, chapter 5). Six basic types of dialogue are cited- persuasion dialogue, negotiation dialogue, information-seeking dialogue, inquiry (or investigative dialogue), deliberation dialogue, and eristic dialogue. Each type of dialogue is a normative structure that represents a conventional kind goal-directed conversation in which arguments are put forward for some purpose by two parties. The central type of dialogue analyzed formally in Walton and Krabbe (1995) is the persuasion dialogue. In this type of dialogue, the proponent has a thesis -a proposition she is supposed to prove-and the respondent has a thesis that is the opposite (negation) of that of the proponent. Or in asymmetrical persuasion dialogue, the proponent has a thesis, and the goal of the respondent is to critically question that thesis, raising sufficient doubt so that the proponent fails to prove it (using the appropriate types of arguments, and by the appropriate standard of burden of proof).

As shown in Walton and Krabbe (1995, pp. 127-132), each side uses chains of inferences connected together in a sequence to form a path of argumentation. Ideally, the path of argument put forward by a proponent should move toward her thesis to be proven. Failure to achieve this ideal is sometimes called the "fallacy of irrelevant conclusion" (*ignoratio elenchi*). At any rate, this notion of a path of argumentation in a persuasion dialogue, where the path is aimed towards the arguer's thesis to be proved, is modelled in the formal structure of persuasion dialogue as a connected sequence of inferences of the kind familiar in logic.

What hasn't been modelled yet is the idea of how far along this ideal path an argument used in a particular case has gone. But it could be possible to approximate such a measure contextually by comparing the actual chain of argumentation used in a given case with the ideal path of argumentation required by the framework of the persuasion dialogue appropriate for the case. This comparison is the basis for measuring how heavily a question (or an asserted proposition, or a term) is loaded in a given case. The method of determining how heavily a given question (proposition, term) is loaded is to use the normative model of dialogue to apply to the given details of the actual case, locating the stage in the dialogue where the question was used, and then extrapolating forward from that point to the thesis to be proven, according to the requirements of the normative model. Thus the method combines the actual details of the argumentation given in the text of discourse of a particular case with the normative structure of the model of dialogue.

The empirical method of determining how heavily a term in a question is loaded is to ask the question in a statistical poll, then replace the term with a descriptively equivalent but less loaded, or more neutral term, and then ask the revised question in a poll with a group of respondents selected in the same way as the first group. Essentially, the method is the experimental method used by Schuman and Presser (1984) for determining response effects of question

wording. A simple example is given by Moore (1992, pp. 343-344). A 1985 survey asked respondents whether too little money was being spent on welfare. 19 percent of respondents said 'yes.' But then when a group of respondents selected by the same criteria were asked the same question with the word 'welfare' replaced by the descriptively equivalent phrase 'assistance to the poor,' 63 percent said 'yes.' The difference of 44 points is the so-called "response effect" of the wording in the question. The proposal I am making is that the response effect can be used an indicator of how heavily a term used in question in a given case is loaded. The heaviness measured is one due to the connotations that a term like 'welfare' has for a mass audience, or wide population of speakers of a natural language. For this group of respondents then, the response effect test gives an indication of how heavily the term is loaded in the question, in relation to a reference term that is equivalent (in denotation, but not connotation for that respondent group).

This type of response effect test, it should be stated, is not an indicator of how heavily a proposition or question is loaded, in an absolute sense. But Schuman and Presser (1981, chapter 9) have also experimented with questions that can be used in polls to try to determine whether and how strongly respondents are committed to particular viewpoints, by studying factors like intensity and centrality of passionate attitudes of respondents in polls. In studying how to conduct statistical polls on the controversial issue of gun permits, for example, Schuman and Presser (1981, p. 24) found a tiny group of respondents who rate the issue of gun permits as being of top importance. Once such a group of respondents has been identified, then a context or normative framework of dialogue is also identified. There is a small group of respondents who see the issue of gun permits as being of top importance. On one side are the proponents of gun permits, and on the other side are those who are strongly against requiring gun permits. So if you ask any question relating to this conflict of opinions-for example, a question on how far governments should regulate private conduct-then that question will be loaded, in opposite directions, for these two groups of respondents. So whether a question (term, proposition) is loaded is relative to (a) a particular group of respondents, and (b) to some issue, or framework of dialogue where this group of respondents has a particular thesis or viewpoint that they are strongly committed to.

At any rate, without going into further details here, it seems that some of these experimental techniques used by Schuman and Presser could be used to get some empirical indication of how heavily a proposition or question is loaded for a specifiable group of respondents. Certainly it would seem to be quite possible to get a reasonable empirical estimate of how heavily a proposition or question is loaded for a target group of respondents by putting questions to this group of respondents that test out their pro or contra attitudes towards an issue. But cautionary remarks are in order. You can't just ask a respondent directly what he thinks about something, say, gun licensing,

and then determine whether or not a question or proposition is loaded with reference to this expressed attitude or commitment. For the definition of the term "loaded" we need to work with here has a rationality component built in. So the kinds of questions the empirical researcher needs to ask are subtle. She must make assumptions about what the issue is supposed to be (that defines the thesis the respondent is supposed to be supporting), and then try to test by indirect questions how the given question or proposition is related by a chain of reasoning to that proposition.

4. Evaluating Cases

In evaluating cases of loaded questions of the kind criticized in the logic textbooks, it is generally not necessary to try to measure exactly how heavily the question is loaded. For one thing, the examples cited are often brief, like the classic spouse abuse case, and not enough context is given so that it can be determined whether the question was really loaded in the (supposedly) real case presented, or how heavily it was loaded. The context of dialogue given may simply not provide enough information to definitively prove such determinations. But we could still make presumptions about whether the question was loaded, depending on the context of use that one might impute to such a question, and then reach a conditional evaluation based on these assumptions. Even this conditional type of evaluation can be instructive for students of critical thinking. Nor is it often necessary to run experiments with question wordings to determine response effects, before questioning a question on the grounds that it contains loaded terms. But still, it is comforting to know that the heaviness of the loading of a question can be measured in principle, and that therefore, as exponents of critical thinking, we are not asking students to evaluate something that is totally subjective, in the sense that it cannot be tested or estimated by objective standards.

The examples cited in logic textbooks are frequently cases where some allegation of unethical conduct is made by the question. For example, Copi (1982, p. 110) cites the following question.

CASE 4: DID YOUR SALES INCREASE AS A RESULT OF YOUR MISLEADING ADVERTISING?

In this case, the profiles of dialogue method can be applied, as shown by Copi's follow-up move in the dialogue, cited as the next reply in the sequence, after the respondent answers 'no' to the question: "So you admit that your advertising was misleading. Do you know that your unethical conduct can get you into trouble?" This latter question in the sequence indicates that the proponent is making an allegation of unethical conduct in asking the original question. Much the same remarks are applicable to the spouse abuse question (case 1). So a normative framework of dialogue for case 4 can reasonably be presumed to apply to the case. The context of case 1 indicates that the proponent is attacking the advertising of the respondent as being deceptive,

i.e. unethical (and-or possibly illegal). The respondent is put on the defensive by the question, and must react to it by defending his side of the contested issue. Thus the question is loaded against the respondent's side, precisely because (according to the normative analysis) it contains a thesis advocated by the questioner, and an opposed thesis that needs to be argued for by the questioner, and argued against by the respondent (in relation to the presumed issue at stake).

The context of dialogue in case 4 is not so explicitly that of a persuasion dialogue based on a known issue where there is a conflict of opinions, as it was in case 3, where the issue was that of abortion. But the allegation of ethical impropriety, made in case 4 through the use of the loaded term 'misleading,' implies a context of persuasion dialogue by aiming a path of argument against the side of the respondent. The respondent is made to appear guilty of some breach of ethical conduct by the question, and therefore a kind of presumption is created that the respondent has been put on the defensive, and that to correct the imbalance, he must reply, and try to make some kind of case on his behalf. If he wants to maintains that he is not guilty of the charge made in the question, instead of directly answering the question, he must try to question or even rebut the allegation it makes.

In another kind of case, the question does not necessarily make an allegation that the respondent has done something unethical, but it contains terms that have an emotional connotation in common usage. The following example of the fallacy of complex question is given by Engel (1976, p. 83).

CASE 5 : WHAT ARE YOUR VIEWS ON THE TOKEN EFFORT MADE BY THE GOVERNMENT TO DEAL WITH THIS MONSTROUS OIL CRISIS?

In this case, the adjectives 'token' and 'monstrous' have negative connotations that put a particular spin on the question. This particular type of tactic is frequently used in political debates, for example, in the kinds of questions so often used in the Open Question Period in parliaments and legislatures, of the kind studied in Walton (1989). Once again, as in case 4, a kind of conflict of opinions between the questioner and respondent is created by this spin on the question. If the respondent doesn't agree that the effort is 'token', or that the crisis is 'monstrous,' he will have to make some sort of attack on the question, in lieu of answering it directly. In a political debate, the use of a loaded question makes the question function like an argument that puts pressure on the respondent's side to respond with some sort of defensive argument.

In cases 4, very little, and in case 5, no information is given in the logic textbook about the wider context in which the question was supposedly used in some sort of dialogue exchange. Still, it is not too hard for the textbook user to get the general idea of how such a question is an effective tactic that can be used unfairly or deceptively to get the best of a speech partner in argumentation. Because questions of this sort are so common, and familiar to us from our experiences in everyday argumentation, we quickly get the idea of the sort of

context that might normally be the setting of use for such a question. Still, the textbooks would do better to make clearer how the evaluation of such questions is best conducted in a pragmatic reconstruction of how the question was being used in a context of dialogue appropriate for the given case (on a basis of comparison with how it ought to be used). The two most important analytical tools that should be deployed for this purpose are the profile of dialogue and the method presented above for determining whether and how a question is loaded.

Some of the cases of questions that have been experimentally tested for response effects would be quite interesting to use as examples in textbooks designed to improve critical thinking skills. A classic example is the case cited by Clark and Schober (1992, p. 31) concerning the *forbid-allow* response effect, which relates to the difference between the following pair of questions.

CASE 6: Q1. DO YOU THINK THE UNITED STATES SHOULD FORBID PUBLIC SPEECHES AGAINST DEMOCRACY?

Q2. DO YOU THINK THE UNITED STATES SHOULD ALLOW PUBLIC SPEECHES AGAINST DEMOCRACY?

As Clark and Schober comment (p. 31), both questions are to be answered 'yes' or'no,' but in *Q1* 'yes' and'no' means'forbid' and'not forbid,' respectively, whereas in *Q2* 'yes' and'no' mean 'allow' and'not allow,' respectively. *Q1* and *Q2* are descriptively equivalent, because to forbid a speech is not to allow it. But 54 percent of respondents to Q1 said yes,' while 75 percent of respondents to Q1 said 'no' (Clark and Schober, 1992, p. 31). This response effect is a difference of 21 points. So from a point of view of how heavily they are loaded with connotative values, the two questions are not equivalent.

Use of this kind of case to verify the response effect of such a difference of wording in questions - a factor that might presumably be far from apparent to many users of a critical thinking textbook, and might initially be perceived as trivial - helps to illustrate the significance and practical usefulness of critically evaluating questions of the kind studied in informal logic courses. What is shown is that certain kinds of questions can be loaded in a tricky or subtle way that might not be appreciated initially. When a significant response effect can be indicated empirically, it can help to offset the popular opinion of many beginners to critical thinking that logical and verbal matters of question construction and wording are of no real importance, and are a waste of time to consider in any depth.

5. Replies to Questions About the Method

The use of normative models of dialogue to evaluate argumentation use in particular cases is a fairly new development, and it may seem radical or unfamiliar to many readers who are not used to this technique. Therefore it may be helpful to conclude by replying to certain worries and doubts that are frequently raised about the method proposed above.

The first question frequently asked is how we can use the method to distinguish between loaded questions - or statements or terms-and genuine argumentative moves (that also go some distance towards proving the respondent's thesis false). This question is an excellent one, and fully deserves to be answered, but it is useful to observe that it is a loaded question itself, especially in regard to one particular presumption. The question presumes that there is a dichotomy between loaded questions and genuine argumentative moves. That is, it presumes that loaded questions are inherently non-genuine-that is, illegitimate, spurious, or fallacious. But this presumption, as shown in section one above, is a misconception about the method of profiles of dialogue, as applied to particular cases. As shown in section one, while the spouse abuse question (to cite the classic case) is fallacious as used in some cases, there are cases where asking the question "Have you stopped abusing your spouse?" could be perfectly reasonable. In other words, loaded questions are not always fallacious, or non-genuine argumentative moves. Whether such a question has been used reasonably or not in a given case is determined by constricting a profile of dialogue for the case, and then evaluating the prior sequence of questioning in light of the known details of the case and the appropriate normative model of dialogue for the case.

Despite this misleading presumption, however, the question does express a legitimate and important concern, and in response to it, a distinction needs to be made between two tasks. One is the task of determining, in a given case, how heavily a question is loaded in that case. That task is to be carried out by the method proposed above. But that there is also an important task that has been the primary concern of the logic textbooks, and other sources that have been worried about the use of loaded questions in argumentation from a point of view of critical thinking. This task is the job of evaluating loaded questions, by judging in particular cases exactly whether, how and why such questions are fallacious (or not fallacious, if they are not). It's important to emphasize here that the method of determining how loaded a question is, in a given case, is not the sole means of carrying out this other task. It is one very important tool, to be used in conjunction with other tools-and particularly the profiles of dialogue method-for carrying out the task of argument evaluation.

The second source of worry can be expressed in the following questions and assertions. What is the connection between the normative analysis-using the models of dialogue and the projected chains of argumentation used to determine the distance between the given proposition and the respondent's

thesis- and the empirical test? Why even assume that there is any connection here? To assume such a connection is to assume that people's patterns of responses to questions are, at least to some degree, connected to judgments about the ways in which ideal chains of argumentation might proceed. But in fact, responses to poll questions might be much less rational than that. To respond to the worries expressed by this set of questions and assertions, it is best to begin once again by pointing out certain potentially misleading implications in them.

One factor in the method for judging how loaded a question is in a given case is connected to judgments about the ways in which ideal chains of argumentation might proceed. Citing this factor by itself, in relation to people's patterns of responses to poll questions-which by our own admission, in relation to the problem our method is addressed to, namely the use of loaded and tricky questions in polling-makes the method sound hopelessly idealistic, by making it seem to assume, unjustifiably, that poll respondents really think out, or are aware of, chains of logical reasoning in some ideal model of argumentation. Can we assume that poll respondents are all that rational? It would seem to be unjustified (and worrisome) that the method makes any such assumption.

To respond to these worries, some of the same points need to be repeated from the response given above to the prior doubts. Yes, the method does assume that a respondent in a given case is rational in a certain sense, because it applies a normative model of dialogue, and the projecting forward of a chain of reasoning in that model of dialogue, towards the respondent's thesis in the dialogue. Because it is an ideal model, like all models, it does not apply or fit perfectly to the details of any particular case. Thus the evaluation of a loaded question in any particular case will inevitably be based on certain assumptions that may not fit the case exactly. Indeed the evaluation of any particular case will have to be, in certain respects, and to some extent, conditional in nature. Is that a problem? The answer is that it is not, as the evaluation of the six cases in section four showed. What the evaluation of a case does is to place the asking of the question in the given case in relation to a context of dialogue that can be extrapolated (by presumption) from the information given in the text of discourse of the case. For example, in case 4, the issue posed by the asking of the question was inferred from the allegation made in the question that the respondent was guilty of misleading advertising. So the question is evaluated as loaded from the point of view that one can reasonably attribute to the respondent (in the absence of indications to the contrary). The thesis that one can reasonably presume that the respondent has an interest in defending as his thesis is the negative proposition that his advertising is not "misleading."

Another point to be made about this particular method of judging loaded questions is that it is tailored to the data given in the particulars of the individual case it is being applied to. What is especially important to note is that the evaluation of how heavily a question is loaded in a given case depends

on the commitments of the respondent, insofar as these are known in the given case. And how heavily a question is loaded depends on that respondent, or on a group of respondents, in the case of a statistical poll. So it is important to realize that, according to the method proposed above, a question can be heavily loaded for one respondent, and less heavily loaded (or not loaded at all) for another respondent. So the determination of how heavily a question is loaded in a given case is not as heavily idealized as the worries expressed above suggest.

Another factor to be clear on is that the test for response effects used by statisticians is far from a perfect instrument. It arises from the worry that a question used in a poll might be loaded or biased in a way that the pollster did not anticipate, and that would make the announcement of the finding of the poll (expressed a scientific finding, with numerical measures of chance of error announced) misleading to users of the poll. Testing for a response effect is one way of getting confirming evidence that a problem of this kind exists in a given case. And the validity of the test, in a given case, depends on statistical assumptions about sampling - in particular, on the assumption that the respondents used in the second test are the same kind of respondents as those used in the first test. These are the usual statistical assumptions about selecting a sample of respondents using methods that do not contain biases of various kinds well known to statisticians. So there are plenty of worries about any particular running of a statistical test for a response effect. It is not a perfect instrument, any more than any statistical test used in any poll is.

6. Summary of the Method

The method proposed for evaluating whether and how heavily a statement, term or question is loaded in a given case is the following. A statement is loaded in a particular case (in a context of dialogue) if it goes some distance towards proving the respondent's thesis false. A term is loaded if it occurs in a loaded statement. A question is loaded if it contains or presupposes terms or propositions that are loaded. So defined then, whether and how heavily a proposition (term, question) is loaded in a given case is determined by (1) the commitments of the respondent in a dialogue, (2) more generally, the context of dialogue for the given case, insofar as that is known, (3) the type of dialogue the participants are supposed to be engaged in (insofar as that is known, or has been determined), and (4) the projected distance between the proposition (term, question), as used in the dialogue exchange, and the respondent's thesis. This method involves the applying of a normative (ideal) model of dialogue to the argumentation used in a particular case, embedded in the text of discourse given in an actual case. How heavily the proposition (term, question) is loaded is determined by applying the normative model of dialogue to the given details of the particular case. In carrying out such an application in a given case, it is necessary to also use the tool of the profile of dialogue, as explained in section

one above. In addition to this normative method, an empirical test is proposed. A question is loaded, or contains loaded terms, if individuals respond to it differently in statistical polls when it is reformulated in an equivalent way using other words. It is not necessary, or practically useful, to use this empirical test to evaluate how heavily a question is loaded in all cases (of the kind illustrated above as typically of concern in logic and critical thinking textbooks). Still, in principle, it is useful to see that there is an empirical test that can be used for this purpose, and that is very useful particularly in evaluating questions used in statistical polls.

What an empirical test for a response effect tells you is that there is something in wording of the question - either in the connotations of the words used, or in the logical structure of the question, generally, that is loading or biasing the question in a certain direction, in relation to the conventional word usage and commonly accepted opinions of a particular class of respondents, chosen by various criteria of the kinds commonly used or recommended by statisticians for polling. The test is a clue or indicator of a trick or twist in the wording of the question that a pollster may not have been aware of, until it was revealed and confirmed by the test. This empirical test does not measure how heavily a question is loaded, generally. It is only designed to measure unanticipated response effects due to question wording. So it measures how one term is loaded in a question, in relation to how another (apparently equivalent) term is loaded in the same question, in relation to a selected group of respondents. Thus it tell us something about loaded terms, but it does not measure, nor has anyone claimed that it measures, how heavily a question is loaded with respect to a particular respondent, or group of respondents.

So why should we assume that there is any connection between this kind of empirical test for response effects of the wording of a question, and the normative and pragmatic method of evaluating how heavily a question is loaded in a given case that was proposed above? The reason is that the pragmatic method proposed above judges how heavily a question is loaded, in a given case, in relation to the commitments of the respondent in that actual case (insofar as these are known, or can be judged by applying the profile of dialogue to the particulars of the case). So when the statistical test for response effects of the wording of a question is run, on a particular group of respondents, where the characteristics of this group has been chosen by the statistical survey methodology, what is being estimated is how heavily the question is loaded, in virtue of its wording, or key terms used in the question, for that group of respondents. So when it is judged that the question is loaded or 'biased' to a certain degree, for that particular group of respondents (as extrapolated to a wider population by the usual methods of statistical surveying), the result of the test is the finding that the question is loaded, in virtue of some term contained in it, for that group of respondents. In using such a test, it is both natural and necessary to make assumptions, and use these assumptions about the known views or positions of this group of respondents.

These factors, which correspond to the known characteristics of the group selected, represent (in the normative model) the thesis (and the known commitments generally) of the respondent.

7. Epistemological Implications

Collection of data in research and in statistical polls and public opinion surveys is always based on the asking of questions. But if natural language wording is used in the question, the question is bound to be more or less loaded in a particular direction, and therefore bound to have significant persuasive implications. Yet the way the results of scientific investigations and surveys are typically announced, it appears the finding has been obtained by an objective and unbiased collection of data. The epistemological significance of this dissonance between appearance and reality should be regarded as extremely important from a viewpoint of critical thinking. We need to realize that before we act on these omnipresent polls, or take them seriously as results that have been scientifically validated and proved, and as constituting good evidence for a conclusion, some way of measuring the bias of the question has to be taken into account. But how do you measure the connotations of words, and the innuendo or implicatures in a question? This is the problem, and statisticians are not used to dealing with this kind of problem. And indeed, they are not well equipped to deal with it, because the problem is a dialectical one, best solved by pragmatic methods on a case by case basis.

Much more work needs to be done by social scientists to investigate the relationship between logical models of dialogue in which argumentation is evaluated as reasonable insofar as it contributes to the goal of the dialogue or not, and empirical methods of judging the attitudes and commitments of respondent groups in statistical polls and surveys. Normative structures of dialogue that have been developed (as cited above) are especially interesting precisely because they do seem to model everyday argumentation much more naturally than the deductive and inductive logical calculi that have dominated the field of logic so heavily (even exclusively) in the past. But exactly how the formal dialectical modeling of loaded questions (and the tricky problems associated with them) correlates with statistical techniques for empirically estimating attitudes and commitments of respondents is really a field of inquiry in its own right that, so far, has been very little explored. As more and more people become aware of the abuses of statistics in push-polling and similarly dubious uses of questioning so widely operative now in the collection of data, the need for taking critical thinking seriously in this area has become more and more evident.

References

Jeremy Bentham, *The Book of Fallacies*, in vol. 2 of *The Works of Jeremy Bentham*, ed. John Bowring, New York, Russell and Russell, Inc., 1962, 375-487 (first published in 1838).

Herbert H. Clark andMichael F. Schober,'Asking Questions and Influencing Answers,' in *Questions About Questions : Inquiries into the Cognitive Bases of Surveys*, ed. Judith M. Tanur, New York, Russell Sage Foundation, 1992, pp. 15-48.

Irving M. Copi, *Introduction to Logic*, 6th ed., New York, Macmillan, 1982.

Cynthia Crossen, *Tainted Truth : The Manipulation of Fact in America*, New York, Simon and Schuster, 1994.

Diogenes Laertius, *Lives of Eminent Pbilosopbers*, vol. 1, Loeb Classical Library, Cambridge, Mass., Harvard University Press, 1959.

S. Morris Engel, W*itb Good Reason: An Introduction to informal Fallacies*, New York, St. Martin's Press, 1976.

Charles Hamblin, *Fallacies*, London, Methuen, 1970. Reprinted by Vale Press, 1986.

Charles Hamblin,'Mathematical Models of Dialogue;' *Theoria*, 37, 1971, 130-155.

David Harrah, The Logic of Questions, in *Handbook of Pbilosopbical Logic*, vol. 2, ed. Dov Gabbay and F. Guenther, Dordrecht, Reidel, 1984, 715-764.

Jaakko Hintikka, 'Semantics and Pragmatics for Why-Questions,' *The Journal of Pbilosophy*, 92, 1995, 636-657.

William and Martha Kneale, *The Development of Logic*, Oxford, Clarendon Press, 1962.

Erik C. W. Krabbe,'So What? Profiles for Relevance Criticism in Persuasion Dialogues,' *Argumentation*, 6, 1992, 271-283.

Jim Mackenzie, The Dialectics of Logic, *Logique et Analyse*, 94, 1981, 159-177.

Michel Meyer, *Problematology: Pbilosopby, Science and Language*, Chicago, University of Chicago Press, 1995.

David W. Moore, *The Super Pollsters*, New York, Four Walls Eight Windows, 1992.

Howard Schuman and Stanley Presser, Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Context, New York, Academic Press, 1981.

Frans H. van Eemeren and Rob Grootendorst, *Argumentation, Communication and Fallacies*, Hillsdale, NJ., Erlbaum, 1992.

Douglas N. Walton, *Question-Reply Argumentation*, New York, Greenwood Press, 1989.

Douglas N. Walton, *APragmaticTheory of Fallacy*, Tuscaloosa, University of Alabama Press, 1995.

Douglas N. Walton and Erik C. W. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, Albany, State University of New York Press, 1995.

Notes

Requirements for the use of non-sexist language are met by following the convention of always making the proponent in a dialogue 'she' and the respondent 'he.'

I would like to thank the Social Sciences and Humanities Research Council of Canada for a research grant that funded this work, and the University of Winnipeg for granting me study leave. The work for the paper was carried out while I was a Research Associate in the Department of Philosophy at the University of Western Australia, and subsequently, Distinguished Visiting Research Fellow the Oregon Humanities Center (University of Oregon). I would like to thank both groups for providing facilities and support.