Understanding Scientific Inquiry

Peter J. Veazie[1]

Abstract

Science is a process of inquiry: a process of asking and answering questions. However, a good question is more than an interrogatory, and a good answer is more than information: there are logical constraints that dictate when a question is answerable and what qualifies as an answer. This paper will provide an understanding of (1) when a question is answerable, (2) when a question is not ready to be asked, (3) when a question is trivial, (4) what is required for a response to be an answer, and (5) what sequence of inquiry is required to identify an answer. Equipped with this understanding, a scientist can better determine an appropriate sequence of study for a research program as well as identify the necessary arguments to warrant claims of understanding, funding, and the publication of research findings.

Keywords: Erotetics; research questions; research philosophy.

---
[1] University of Rochester, Rochester New York, USA
peter_veazie@urmc.rochester.edu

1 Introduction

A scientist answers questions. To do this well requires an understanding of what is a good question and a good answer, which in turn requires understanding the logic of inquiry. This paper will explain key properties of questions and answers that constrain research and guide research programs thereby clarifying requirements for questions, answers, necessary preliminary work, and possible future research. Note, this article regards applied philosophical issues, it does not address the psychology or empirical requirements of scientific inquiry.

There are various approaches to the logic of questions and answers in the linguistics and philosophy literatures (formally called erotetics). For information on some approaches see the references by Scha and Wiśniewski (Scha, 1983, Wiśniewski, 1995). However, Rescher's book Inquiry Dynamics (Rescher, 2000), which underlies this presentation, may be more directly applicable.

2 The basics of questions and answers

Questions are requests for specific information. If I ask `Can stress cause depression?′ and receive the response `Stress can cause high blood pressure′, this response may be informative but it does not satisfy the question. If a response does not satisfy the question, then it is not a possible answer to that question. A response satisfies a question, right or wrong, if given it were understood it would not be reasonable to repeat the question. Consider the question `Is it raining outside?′ being followed by the response `Yes, it is raining outside′, it would not be reasonable to repeat the question: we would not say `But is it raining outside?′. The response satisfied the question. Alternatively, consider if the question `Is it raining outside?′ was responded to by `Green is my favourite color′, it would be reasonable to say `But is it raining outside?′. If the response to `Can stress cause depression?′ is `Stress can cause high blood pressure′, it would be reasonable to reply with `But can stress cause depression?′. Assuming the response is understood, the need to repeat the question is an indicator that the response was not a possible answer. A possible answer need not be correct to satisfy a question in this sense.

Although an answer resolves the question-specific inquiry, not all questions are answerable: those that presuppose propositions taken to be false are not answerable. Consider the question `Has Fred stopped petting his dog?′. This question presupposes, among other things, that Fred has or had a dog. If this is taken to be false (i.e. Fred is believed to never have had a dog), then the question is not answerable. Neither yes nor no suffices. It would not make

sense to say `Fred does not have nor has ever had a dog, but has Fred stopped petting his dog?′. To be answerable there must be at least one response that satisfies the question. Here is another example: `How does stress cause depression?′. This question presupposes that stress causes depression; if the proposition *stress causes depression* is taken to be false, then the question of how it does so is not answerable. It would be nonsensical to say `Stress does not cause depression, but tell me how stress causes depression?′ A question that is not answerable is an *illegitimate* question.

If, on the other hand, the truth of a question's presuppositions is unknown (I do not know whether Fred has or ever has had a dog or whether stress causes depression), then the legitimacy of the question is unknown. In this case the question is *premature*. Premature questions lead to new questions regarding the unknown presuppositions that must be answered before legitimacy can be determined. If the truth of the proposition *stress causes depression* is unknown, then the question `How does stress cause depression?′, which presupposes this proposition, is premature. Legitimacy depends first on answering the question `Does stress cause depression?′ If the answer is `Yes, stress causes depression′, then the question regarding *how* stress causes depression is legitimate. If the answer is `No, stress does not cause depression′, then the *how* question is illegitimate.

A question that is not an inquiry is *trivial*. A trivial question is one for which the answer is taken to be known. Consequently, a question is trivial once it is determined there is only one possible answer since this possible answer must be *the* answer; no inquiry is required. Also, a question is trivial once the actual answer from a complete set of possible answers is taken to be determined; again, no further inquiry is required. To claim a question is trivial, however, is not to claim its answer is unimportant, it is only to say that since we presume to know the answer, the question itself does not require inquiry.

It should be clear that inquiry leads to a sequence of questions that can be tracked backward or forward. A premature question requires answers to questions regarding its presuppositions. This type of reasoning allows the scientist to identify work that must be completed before the target question is legitimate. Since the answer to a question can be the presupposition for other questions, inquiry is a forward-oriented process as well. If the question `How does stress cause depression?′ is legitimate and we take its answer to be `Stress causes depression via mechanism M1.′, then based on this answer we might ask `How does M1 cause depression?' or `Why does stress cause depression by M1 rather than mechanism M2?′ (which also presupposes M2 does not explain). And so, our research program regarding stress and depression continues.

Responses also have presuppositions. For example, the question `Does stress cause depression?′ has a response of `Yes, stress causes depression′, but this response cannot be true unless both stress and depression exist. If one of these presuppositions is taken to be false, then the response is considered false and thereby cannot be an answer to the question. The response `No, stress does not cause depression′, however, does not presuppose stress and depression exist; indeed, this response may be offered for the very reason that stress or depression are taken not to exist. Consequently, if the presuppositions of each response are taken to be true, the set of possible answers to the question `Does stress cause depression?′ is {`Yes, stress causes depression′; `No, stress does not cause depression′}. However, if the presupposition that depression exists is taken to be false, then the set of possible answers is {`No, stress does not cause depression′}, which comprises only one response and is thereby trivial—we have what we take to be the answer.

Propositions can also presuppose the truth of other propositions. For example, the proposition *stress causes depression* is true only if the propositions *stress exists* and *depression exists* are true; consequently the first proposition presupposes the other two. Presuppositions are also transitive: if one statement or proposition, say *s*, presupposes another proposition *p*, and the proposition *p* presupposes the proposition *q*, then *s* also presupposes *q*. For example, the question `How does stress cause depression?′ presupposes *stress causes depression*, and *stress causes depression* presupposes *stress exists* and *depression exists*; therefore, the initial question not only presupposes *stress causes depression* but also *stress exists* and *depression exists*. This transitivity generates a hierarchy of presuppositions. In the preceding example, the initial question is premature until the proposition *stress causes depression* is confirmed. This requires answering the implied question `Does stress cause depression?′, but this question is itself premature until its presuppositions are established. So, before these questions are legitimate, the questions `Does stress exist?′ and `Does depression exist?′ must be affirmatively answered. A legitimate line of inquiry must respect the presupposition hierarchy.

For notational purposes, I will indicate a set of responses by $\{_R \dots \}$, and set of possible answers by $\{_A \dots \}$, and the list of presuppositions by $<\dots>$ (however, I will not include all presuppositions, only those being discussed). A question Q that has a response set with k responses $\{_R r_1, r_2,..r_k\}$ may be written as $Q\{_R r_1, r_2,..r_k\}$. A question Q that presupposes a proposition p is indicated as $Q<p>$. A response r that presupposes proposition p is indicated as $r<p>$. And, a proposition p that presupposes q is written as $p<q>$. We can nest the presupposition delimiter to represent a presupposition hierarchy. For example, we can denote the case in which Q has response r that presupposes p, and p presupposes q, as $Q\{_R r<p<q>>\}$; moreover, for greater clarity we can

provide subscripts to the presupposition delimiters to facilitate tracking the nested lists: for example, Q{$_R$ r<$_1$ p_a <$_2$ q_a, q_b $_2$>, p_b $_1$>} indicates presupposition list 1 includes p_a and p_b, and presupposition list 2 (which are presuppositions of p_a) includes q_a and q_b. Although r<p<q>> implies r<p, q>, I will use the former notation to represent the latter. In other words, I will leave the implication of the transitive property of presuppositions implicit: writing that *r* presupposes *p*, which presupposes *q*, as r<p<q>> will be taken to imply that *r* also presupposes *q*. If I wish to denote a question, its answer set, and the question's presuppositions but not response-specific presuppositions, it would be indicated as Q{$_A$ r_1, r_2,..r_k}<p<q>>. I indicate all presuppositions at the answer level as Q{$_A$ r_1<p, p_1>,r_2<p, p_2>,..r_k <p, p_k>}, noting that *p*, being common across all answers is therefore a presupposition for the question itself, whereas the p_i denote response-specific presuppositions. The question-specific presuppositions can be "factored" out of the last notation so that Q{$_A$ r_1<p_1>, r_2<p_2>,..r_k<p_k>}<p> means the same as Q{$_A$ r_1<p, p_1>, r_2<p, p_2>,..r_k<p, p_k>}.

A presupposition of a question, as opposed to a presupposition of a response, is a proposition that all possible answers presuppose. Consider the example `How does stress cause depression?′ {$_R$ `Stress causes depression via a cortisol mechanism′<stress causes depression, cortisol exists>; `Stress causes depression via an anxiety mechanism′<stress causes depression, anxiety exists>}. Each response presupposes the existence of a different mechanism (cortisol or anxiety), and both responses presuppose stress causes depression. There is no way to answer the question without presupposing the common proposition that stress causes depression; therefore, if this proposition is false, the question cannot be answered and is thereby illegitimate. The question and its presupposition are `How does stress cause depression?′<stress causes depression>, which also presupposes that stress exists and depression exists. If a scientist submits an application for funding proposing to answer the question `How does stress cause depression?′ but has not established *that* stress causes depression, then reviewers may deem the question to be premature and request the proposition *stress causes depression* be established first.

What happens if a response-specific presupposition is taken to be false? That response cannot be a possible answer and should be excluded from the set of possible answers: stress cannot possibly cause depression via anxiety if anxiety does not exist. Consider the following example and whether the response set is also the set of possible answers.

> `Is it raining outside?′ {$_R$ `Yes, it is raining outside′; `No, it is not raining outside′}

> To determine the set of possible answers we need to consider each response and its presuppositions. The response `Yes, it is raining outside′

presupposes *rain exists*, *outside exists*, and *rain is the type of thing that can occur outside*. `No, it is not raining outside′ has no such presuppositions. So, indicating the presupposed propositions related to each response gives

> `Is it raining outside?′{$_R$ `Yes, it is raining outside′ <*rain exists*, *outside exists*, *rain can occur outside*>; `No, it is not raining outside′ <>}

If all presuppositions are taken to be true, then each response is a possible answer and the set of possible answers is the same as the indicated response set:

> `Is it raining outside?′{$_A$ `Yes, it is raining outside′; `No, it is not raining outside′}.

Does the set of possible answers to all `Is X the case?′-type questions always contain both Yes and No? Consider the following example.

> `Is there a Unicorn in the kitchen?′{$_R$ `Yes, there is a unicorn in the kitchen′; `No, there is not a unicorn in the kitchen′}

The response `Yes, there is a unicorn in the kitchen′ presupposes unicorns and kitchens exist as well as unicorns are the type of thing that can be in kitchens. `No, there is not a unicorn in the kitchen′ has no such presuppositions. So, including the presupposed propositions gives:

> `Is there a Unicorn in the kitchen?′{$_R$ `Yes, there is a unicorn in the kitchen′ <*unicorns exist, kitchens exist, unicorns can be in kitchens*>; `No, there is not a unicorn in the kitchen′ <>}

If the presupposition that unicorns exist is not taken to be true, then the `Yes…′ response is not a possible answer and the set of possible answers does not include this response:

> `Is there a Unicorn in the kitchen?′{$_A$ `No, there is not a unicorn in the kitchen′}

This is a case in which the set of possible answers contains only one response. A question with a singleton as a set of possible answers is trivial, which is to say when we have only one possible answer, we have the answer; no further inquiry is required (i.e. there is no need to look for unicorns in the kitchen). Ultimately, the goal of scientific inquiry is typically to identify the actual answers, but if we start with only one possible answer, then there is no need for inquiry.

3 The logic of inquiry applies to beliefs

I have been stating whether propositions have been taken to be true or taken to be false rather than stating whether they are in fact true or in fact false. This is because the logic of inquiry is dependent on a given set of beliefs. Consequently, a question may be legitimate, premature, or trivial according to one set of beliefs and not so according to another; and, a response may be a possible answer in one set of beliefs and not a possible answer in another. Therefore legitimacy, prematurity, triviality, and the status of responses can differ across people. If Lisa believes depression exists, Fred believes that depression does not exist, and I am unsure then (assuming all other presuppositions are believed by all) the question regarding how stress causes depression is legitimate for Lisa, illegitimate for Fred, and premature for me. If Lisa were to pose the question to Fred and me, Fred may well say `Your question presupposes something that does not exist and is therefore not answerable′, and I may reply `I cannot answer your question because I do not know whether it is answerable.′ Of course, Lisa may seek answers for herself since to her the question is legitimate. If all three of us were reviewing an application for funding to address the question, then Lisa may be willing to consider the substantive and methodological merits of the proposal (it is after all a legitimate question), Fred would reject the proposal out of hand (no point in considering the merit of research to answer an unanswerable question), and I would recommend that the question of depression's existence must be answered first (no point in considering the merit of research to answer a question that may not be answerable).

4 From possible answers to actual answers

Having a legitimate question and corresponding set of possible answers is necessary but not sufficient for successful inquiry. Successful inquiry requires culling what we take to be *the* answer from the set of possible answers. Understanding this step requires understanding why in the preceding sections answers are demarked with single quotes but propositions are indicated using italics and no quotes. What is being differentiated here is a statement from its content, an assertion from the content of that assertion. For example, `There is a unicorn in the kitchen′ is a statement that asserts the proposition *there is a unicorn in the kitchen*; `The kitchen has a unicorn in it′ is another statement that asserts the same proposition that *there is a unicorn in the kitchen*. It is in virtue of this distinction between an assertion and its content that we can reduce a set of possible answers to the set of actual answers by addressing the hierarchy of presuppositions.

An assertion always presupposes its content. Not only does the assertion `Stress causes depression′ presuppose *stress exists* and *depression exists*, but also that *stress causes depression*. In fact, the assertion is first and foremost presupposing its content, and then the presupposed content itself presupposes more elementary propositions. So, a more accurate representation of the structure for this example is `Stress causes depression′ *<stress causes depression <stress exists, depression exists>>*. If we apply this hierarchical structure to a question we get, for example:

> `Does stress cause depression?′ {$_R$`Yes, stress causes depression′*<stress causes depression <stress exists, depression exists>>*; `No, stress does not cause depression′*<stress does not cause depression>*}

This hierarchy dictates the necessary sequence of inquiry. The response `Yes, stress causes depression′ presupposes its content *stress causes depression*, which in turn presupposes that stress and depression exist; consequently, we need to first address the deepest presuppositions that *stress exists* and *depression exists* before we can address the higher-level presupposition of the content proposition *stress causes depression*. Ultimately if we confirm or disconfirm the content proposition, we will in this case have only one answer remaining in the response set. Notice that if `Yes…′ is the answer, then `No…′ cannot be the answer: one answer logically precludes all others, which is not generally true of all sets of possible answers.

Consider the sequence of inquiry required for the question `How does stress cause depression?′, delineated below by indentation for clarity.

> `How does stress cause depression?′
> {$_R$`Stress causes depression via cortisol′
> $<_1$ *stress causes depression via cortisol*
> $<_2$ *stress causes depression*
> $<_3$ *stress exists, depression exists* $_3>$,
>   *cortisol exists* $_2>_1>$;
> `Stress causes depression via anxiety′
> $<_4$ *stress causes depression via anxiety*
> $<_5$ *stress causes depression*
> $<_6$ *stress exists, depression exists* $_6>$,
>   *anxiety exists* $_5> _4>$}

The first response to the question regards cortisol. This response presupposes its content *stress causes depression via cortisol*. Note that the first list of presuppositions for any response always contains only one proposition, the content proposition. In this example, we can address this proposition by answering its related question `Does stress cause depression via

cortisol?′, but this is a premature question if its presuppositions are not already established. The content proposition of the response presupposes *stress causes depression* and *cortisol exists* (presupposition list 2). We can address these presuppositions by asking their related questions `Does stress cause depression?′ and `Does cortisol exist?′. However, the first of these questions, `Does stress cause depression?′, presupposes *stress exists* and *depression exists* (presupposition list 3) and is premature if these propositions are not taken to be true. Therefore, we need to start by establishing whether stress and depression exist, as well as whether cortisol exists before the higher order presuppositions that depend on them can be addressed and ultimately the content proposition, presupposition list 1, can be used to determine whether the response can be taken as an answer to the question. The same logic applies to the second response. Once a complete set of answers is culled from the set of all possible answers, the question becomes trivial.

From this perspective, the set of possible answers are the responses that can resolve the question-specific inquiry and for which all presuppositions of the content propositions are presumed true, leaving the content propositions themselves undetermined. The final answers from this set of possible answers are determined by investigating which of the content propositions are taken to be true. This set of answers need not be a singleton: there may be more than one answer, in which case their conjunction can be considered a single answer.

Do we need the full set of possible answers to engage inquiry into a legitimate question? No. We can focus on whether a given response provides an answer without knowing all possible responses. For example, `How does stress cause depression?′ may have a very large, even unknown, set of possible answers, but we can focus on the response `Stress causes depression via cortisol′ without knowing the other possible answers by asking `Does stress cause depression via cortisol?′. If the answer is yes to this question, then we have an answer to the *how* question. However, this answer is incomplete. Knowing that cortisol answers the *how* question does not mean that there are no other answers as well (perhaps anxiety also causes depression). If instead we take the answer to be that cortisol does not answer the *how* question, this does not tell us what does. So, we can inquire into a question without knowing the complete set of possible answers, but we may not be able to provide a complete answer to the question by doing so.

5 Inquiry as a process of belief change

Inquiry is a process by which beliefs are changed. Any given set of beliefs comprise propositions that can be presupposed in questions. For example, the proposition *stress causes depression*, if taken to be true, can be a presupposition for asking `How does stress cause depression?'. Also, each proposition itself can be reconsidered. For example, the proposition *stress causes depression* can be reconsidered by the question `Does stress cause depression?'. Reconsideration is a process of opening up a previously closed question: going from a set of answers to an expanded set of possible answers: for example, going from `Does stress cause depression?'{*A*`Yes, stress causes depression'}, which is trivial, to `Does stress cause depression?'{*A*`Yes, stress causes depression'; `No, stress does not cause depression'}, which is not trivial.

Consequently, we can think of a set of all legitimate questions allowed by a given belief set as including all questions that the belief set can support as presuppositions having answers that produce new beliefs and questions that reconsider existing beliefs. Together these questions allow for changes in a belief set. The answer to a new legitimate question increases the set of believed propositions (and thereby expands the set of possible legitimate questions), whereas reconsideration can change the existing belief set by substituting the proposition of a new answer for the old answer. This should be clear from the discussion above: if `Does stress cause depression?' is legitimate and not a reconsideration, then its answer will constitute a new proposition taken to be true—the belief set is thereby expanded. If on the other hand the proposition that *stress causes depression* is a current belief, the question `Does stress cause depression?' is a reconsideration, which if it is subsequently answered `No, stress does not cause depression' will constitute a change in belief by replacing *stress causes depression* with *stress does not cause depression*.

A change in belief due to a reconsidered proposition has ripple effects. For example, all propositions in the belief set that presupposed a reconsidered proposition now taken to be false are no longer valid. Any proposition that is recognized as no longer valid is no longer logically in the belief set. For example, suppose we reconsider the belief in the proposition that *stress exists*: `Does stress exist?'{*A* `Yes, stress exists'<*stress exists*>, `No, stress does not exist'<*stress does not exist*>}. Suppose further that the result of our reconsideration is `No, stress does not exist', then not only does the belief set replace *stress exists* with *stress does not exist*, but all beliefs that presuppose *stress exists* are no longer valid.

6 Conclusion

The logic of inquiry presented above provides an understanding of questions and answers to guide scientists in identifying when a question is premature, when a question is legitimate, when a question is trivial, and how to identify when a response is a possible answer and when it is indeed the answer. There are five main points to remember:

1. A question that presupposes a proposition taken to be false is not answerable and is thereby illegitimate. For example, the question `How does Regulatory Focus Theory explain variation in risk tolerance?′ is illegitimate if it is believed that risk tolerance does not vary. Such a question, with a presupposition taken to be false, should not be posed.

2. A question that presupposes a proposition the truth of which is unknown is premature. Belief regarding the truth or falsity of the presupposition must be determined before the question can be identified as legitimate or illegitimate. For example, the question `How does Regulatory Focus Theory explain variation in risk tolerance?′ is premature if variation in risk tolerance is not established. Such a question requires additional work to establish the requisite presupposition before it can legitimately be posed.

3. A response is a possible answer if it can satisfy the question and if all the propositions presupposed by the content of the response are taken to be true. If a presupposition of a response's content proposition is taken to be false, the response is not a possible answer.

4. Actual answers are the subset of possible answers for which their presupposed content is taken to be true.

5. A question is trivial if its set of possible answers contains only one response or only responses with content propositions all taken to be true. For example, the question `Does Regulatory Focus Theory explain age disparities in communication strategies?′ is trivial if the phenomenon of age disparities in communication strategies does not exist. The `Yes…′ response presupposes the phenomenon; therefore, if the phenomenon is taken not to exist, it is not a possible answer. In this case the `No…′ answer is the only possible answer and is therefore *the* answer, taken to be known *a priori*. Consequently, the question is trivial. As another example, consider `How does stress cause depression?′ $\{_A$ `Stress causes depression via cortisol'; `Stress causes depression via anxiety'}. If we take this to be the complete set of actual answers (i.e. we take it that all answers in the set are in fact answers and that there are no other actual answers not included in the set), then the question is trivial. Explanation-seeking questions (why

and how questions) are often never formally trivial because a complete answer is often not achieved.

Understanding the logic of inquiry will help in formulating research programs by identifying the presupposition hierarchy underlying their ultimate goals. This understanding will also assist in writing applications for research funding and writing scientific papers. The background for such an application or paper should provide compelling arguments for all contestable presuppositions lest the question be deemed illegitimate or premature by the reader. Similarly, presuppositions of any proposed answer need to be established before it can be considered a possible answer.

Although the preceding has been framed in terms of the overall research question, the same logic applies to any question or assertion (to assert a proposition is to affirmatively answer the question regarding the proposition's truth). For example, to claim that a particular data set provides evidence for a hypothesis presupposes that the measurement and methods underlying the evidential claim are valid and reliable thereby requiring affirmative answers to methodological questions prior to making the evidential claim. These answers can be essential background for successful funding applications and research papers.

The process of scientific inquiry requires nuanced considerations of credibility and belief. In its actual application, we would say that the degree to which a question is legitimate depends on the degree to which its presuppositions are taken to be true or taken to be false. Similarly, for any proposition, its credibility is dependent upon the credibility of its presuppositions. The better evidence and argument that can be amassed in support of a proposition, the more credible is that proposition and the more confident we can be in using it as a presupposition for questions, assertions, and other propositions.

In general, a careful consideration of the presupposition hierarchy for inquiry provides guidance regarding both the sequence of study underlying a research program as well as highlighting the background required for compelling funding applications and manuscript submissions.

References

RESCHER, N. 2000. *Inquiry dynamics,* New Brunswick, N.J., Transaction Publishers.

SCHA, R. J. H. 1983. *Logical Foundations for Question Answering,* Eindhoven, the Netherlands, Philips Research Laboratories.

WIŚNIEWSKI, A. 1995. *The posing of questions : logical foundations of erotetic inferences,* Dordrecht ; Boston, Kluwer Academic Publishers.