THE 'GREEN CHEESE' AND 'RED HERRING' PROBLEMS RECONSIDERED. EPISTEMOLOGICAL VS. METHODOLOGICAL TASKS FOR PHILOSOPHERS OF SCIENCE

Erik Weber
Raoul Gervais
Jeroen Van Bouwel

Centre for Logic and Philosophy of Science
Ghent University (UGent)
Blandijnberg 2, B-9000 Gent, Belgium
Erik.Weber@ugent.be
Raoul.Gervais@ugent.be
Jeroen.VanBouwel@ugent.be

Abstract
In this paper, we show that the so-called "green cheese" and "red herring" problems, which have been raised in connection to Bas van Fraassen's pragmatic theory of scientific explanation, lead to two theoretical (or epistemological) tasks for philosophers of science and two practical (or methodological) tasks. After clarifying these tasks we argue that the theoretical tasks are extremely difficult. We are pessimistic about their accomplishment and motivate this. The practical/methodological tasks are easier to handle, and we make some substantial contributions with respect to these tasks. Finally, we compare our practical/methodological results to Kareem Khalifa's hands-on contextualism.

Key words
Epistemic interests, erotetic model, explanation, green cheese, red herring.

THE 'GREEN CHEESE' AND 'RED HERRING' PROBLEMS RECONSIDERED. EPISTEMOLOGICAL VS. METHODOLOGICAL TASKS FOR PHILOSOPHERS OF SCIENCE

1. Introduction

In chapter 5 of *The Scientific Image* Bas van Fraassen developed his pragmatic theory of scientific explanation. Van Fraassen claims that explanations are answers to contrastive questions of the form "Why X rather than Y?". Critics of van Fraassen's account claim that he fails to place sufficient restrictions both on what counts as a legitimate why-question and on what counts as a relevant answer to a why-question. These critics point out that his account faces two types of counterexamples: (1) there are why-questions which we intuitively judge as not legitimate but are legitimate why-questions according to the erotetic model; and (2) there are answers that we intuitively see as inadequate explanations but are adequate according to the erotetic model. Mark Risjord (2000) introduced two now common labels: the existence of the first type of counterexamples is now usually referred to as the "green cheese" problem; the existence of the second type of counterexamples is usually referred to as the "red herring" problem (we will explain where these labels come from in Section 2). Kitcher & Salmon 1987 were the first ones to criticise van Fraassen along these lines (they focused on the relevance of answers, so they only pointed at the second problem).

This paper has four aims:

(1) We want to show that *each* of the problems leads to *two* tasks that philosophers of science can set for themselves: a *theoretical* task and a *practical* task. The theoretical tasks are about fundamental epistemological insights, the practical tasks are of a methodological nature (they are about helping scientists to ask the right why-questions and giving adequate explanations). So there are *four* tasks for philosophers of science originating from the two problems.

(2) We want to show that the two theoretical tasks are very difficult: their accomplishment requires that we satisfy two desiderata that pull in opposite directions.

(3) We want to show that the practical tasks are easier to handle, because there is no such tension and they allow for a piecemeal approach.

(4) We want to make some substantial contributions to the two practical tasks.

The structure of this paper is as follows. In Section 2 we present van Fraassen's erotetic model of explanation and clarify the green cheese and red herring problems. We also show how these problems lead to the theoretical and practical tasks (cf. our first aim). In Section 3 we show that the theoretical tasks are very difficult (cf. the second aim) and in this way motivate our pessimism about them (we think that the chance that someone will accomplish them is very low). In Section 4 we deal with the practical task that originates from the green cheese problem, in Section 5 with the practical task that originates from the red herring problem. So Section 4 and 5 taken together deal with our third and fourth aim. In Section 6 we compare the results from Section 4 and 5 with Kareem Khalifa's hands-on contextualism (Khalifa 2004). The motivation for this comparison is that Khalifa's project fits into the practical tasks we will delineate.

2. Green cheese, red herrings and dead Kennedys.

2.1 The pragmatic theory of explanation regards explanations as answers to why-questions (van Fraassen 1980, p. 134). According to van Fraassen, a why-question is always contrastive.[1] The simplest form is "Why X rather than Y"? Such contrastive questions have two important features: the *topic* (in this case X) which is taken to be true and the *foil* (in this case Y) which is taken to be false. Here are some examples:

> Why does John paint a portrait of the Queen, rather than a landscape?
> Why does John rather than Bill paint a portrait of the Queen?

While van Fraassen introduced the term "topic", he used the term "contrast-class" instead of the now common term "foil". The contrast-class is a set of propositions that contains the (true) topic plus at least one false proposition. In this way, he allows for more complex questions where the topic is contrasted with more than one alternative. For our purposes, the simple contrastive why-questions will suffice.

Besides contrast-classes, Van Fraassen introduces a second important concept: relevance relation. In his view, when we ask a why question, we always have a relation R in mind; R is the relation which answers should have with respect to the contrast in the question.

2.2 Kitcher and Salmon start their criticism of Van Fraasen with the observation that his aim and view are ambiguous:

> We find both the title and the view ambiguous. The purpose of the present discussion is to underscore the difference between a theory of the pragmatics of explanation and a pragmatic theory of explanation. (1987, p. 315)

Kitcher & Salmon are right about this ambiguity. Van Fraassen's suggests at several places that his aim is to develop a philosophical theory of explanations (e.g. p. 134 and p. 146). But he never explicitly formulates his view in a definition. Kitcher & Salmon argue that there are serious problems with van Fraassen's views *if we regard them as a theory of explanation*. This qualification is important for us here, because we take the same perspective: we do not want to criticise nor defend van Fraassen. What we are interested in is: what material does his view offer for someone who wants a theory of explanation? What can we extract from his work for such a project? Are the things we can extract adequate? And are they enough or should we add things? This is our perspective on van Fraassen's work.

Van Fraassen does not formulate any further restrictions on why-questions on top of the ones mentioned in 2.1. So the view on what counts as a legitimate why-question which can be formulated based on his work is:

(Q) It is worthwhile to attempt to answer the contrastive question "Why X rather than

[1] We do not agree with this. We think that some, but not all why-questions are contrastive. Nothing in this paper presupposes that all why-questions are contrastive. But we do presuppose that *some* of them are contrastive.

> Y?" if and only if (a) X is true, (b) Y is false and (c) we have a relation R in mind that answers should and can satisfy.[2]

Incidentally, we believe that requirement (b), i.e. that Y is false, is *too strong*, because it excludes interesting questions. We come back to that in Section 3.1. What is important now is that (Q) has been criticised as being *too liberal.* We will discuss this type of criticism below in 2.2.

The view on what counts as an explanation that can be extracted from the work of van Fraassen, can be summarised as follows:

(A) An answer to a contrastive why-question is an adequate explanation if and only if (a) the question is about a true contrast, (b) the answer is true and (c) it stands in the contextually determined relevance relation R to <X,Y>.

By a "true contrast" we mean a contrast of the form "X as opposed to Y" where X is indeed the case. Van Fraassen is very liberal with respect to the relevance relations. He allows it to vary from context to context:

> So scientific explanation is not (pure) science but an application of science. It is a use of science to satisfy certain of our desires; and these desires are quite specific in a specific context, but they are always desires for descriptive information. … The exact content of the desire, and the evaluation of how well it is satisfied, varies from context to context. It is not a single desire, the same in all cases, for a special sort of thing, but rather, in each case, a different desire for something of a quite familiar sort. (1980, p. 156)

Let us call the conjunction of (Q) and (A) the erotetic model of explanation. This erotetic model is what we get when we extract a theory of explanation from van Fraassen's work. This model has been charged of excessive liberalism. The critics claim that the erotetic model allows too much choice both with respect to foil (the "green cheese" problem) and with respect to the relevance relation (the "red herring" problem). Let us now have a closer look at the these objections.

2.3 First, the green cheese problem. Every state of affairs that we might be explanatorily interested in can always be contrasted with countless other states of affairs that are not the case, but the vast majority of these states of affairs lead to questions that (a) most people would regard as strange and (b) scientists would not consider interesting. For instance, the topic "The level of unemployment in Belgium is 7.4 %[3]" might be contrasted with an infinite set of states of affairs which are not the case, but the vast majority of which lead to strange and uninteresting questions. For example, the question "Why is the level of unemployment in Belgium 7.4 %, rather than the moon made of green cheese?" contrasts the topic with a such a foil. Scientists will never ask this question. On the other hand, the question "Why is

---

[2] "can satisfy" means that we think there exists a true answer which stands in relation R to the contrast.

[3] This and the other figures are the unemployment rates for November 2012, as released by the European Commission on the website Eurostat. Retrieved 22-01-2013 from: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&language=en&pcode=teilm020&tableSelection=1&plugin=1

the level of unemployment in Belgium 7.4 %, rather than 10.7 % (which is the average level for the 27 EU member states)?" might very well receive attention. The green cheese problem consists in the fact that (Q) not only allows for foils of the latter kind (the ones that lead to scientifically interesting questions) but also those of the former kind.

Second, the red herring problem. There are countless relevance relations to choose from, but only a small portion of those single out answers we intuitively consider to be interesting explanations. If any relevance relation is allowed, then even if an adequate foil has been chosen, any true statement can explain any true contrast. Suppose the relevance relation is specified as "the answer must have more *r*s in it than the question". Then the question (to stick to our introductory example): "Why is the level of unemployment in Belgium 7.4 %, rather than 10.7 %?" can be answered by the statement "Because red herrings have gills". Or – to borrow the original example of Kitcher & Salmon 1987 – suppose we ask why John F. Kennedy died on November 22, 1963, rather than November 23. If there is no restriction on the relevance relation, then we cannot exclude statements based on astrological theory as answers: a true description of the positions of stars and planets at the time of Kennedy's birth could count as a relevant answer for the question why Kennedy died on November 22, 1963 rather than on the 23rd. Clearly we want to exclude such inappropriate answers as scientific explanations (because scientists will never propose such answers). The red herring problem then, is how to exclude relevance relations that single out answers to why-question that, though true, are clearly not explanatorily interesting, because they are simply irrelevant.[4]

2.4 Summing up, the erotetic model of explanation suffers from excessive liberalism, both with respect to what counts as an appropriate question, and what counts as a relevant answer. This is not necessarily a criticism of van Fraassen (given that it is not clear what his aims are) but certainly has implications for philosophers who want to use van Fraassen's ideas in order to theorise about scientific explanation. What can such philosophers of science do once they see this?

The theoretical tasks which philosophers of science can set themselves in reaction to this situation are in fact attempts to do better than (Q) and (A). The first possible theoretical task is to fill in the following scheme in a better way than (Q):

> It is worthwhile to attempt to answer the contrastive question "Why X rather than Y?" if and only if […] .

---

[4] It should be noted that some philosophers have construed incorrect examples of the red herring problem. Thus, it has been argued that if the relevance relation states that the answer must have three *r*s and two *g*s in it, then the answer "Red herrings have gills" explains why Bush is president (Khalifa 2004). This however is a wrong example, for 'having three *r*s and two *g*s' is a property of the explanans, *not* a relation between explanans and explanandum. The confusion stems, we think, from of a shift in terminology: Khalifa follows Risjord's substitution of the term 'relevance relation' with 'relevance criterion' (Risjord 2000), and of course, though not a relation, 'having three *r*s and two *g*s' can be a criterion. Although this is not a serious problem (as we have just shown, the relevance criterion can easily by reformulated into "the answer must have more *r*s in it than the question"), nevertheless we feel that to avoid such confusions, it is best to stick to van Fraassen's original term; especially since the terminological shift was not made on principal grounds (Risjord 2000, p. 71 note 4).

The green cheese problem means that (Q), as a way of filling in the "[…]" in the scheme is not good. So we can try to do better. If we manage to accomplish this first task in an adequate way, we obtain an important epistemological insight: we know what makes a contrastive question a real explanation-seeking question.

Analogously, the red herring problem leads to the possible theoretical task of giving a better implementation (compared to (A)) of the following scheme:

An answer to a contrastive why-question is an adequate explanation if and only if […]

If we manage to accomplish this second theoretical task, we obtain another important epistemological insight, viz. what makes an answer to a contrastive question an explanation.

The practical tasks which philosophers of science can set themselves are (1) try to formulate guidelines which help scientists to avoid uninteresting ("green-cheese-like") questions and (2) try to formulate guidelines which help scientists to avoid uninteresting ("red-herring-like") answers. While the theoretical tasks are of a descriptive nature, the practical tasks are methodological: the aim is to offer strategies which help scientists to avoid asking uninteresting questions and considering irrelevant answers to be adequate explanations. We will further clarify the distinction between the two types of tasks at the end of Section 4.

3. The difficulty of the theoretical tasks

In this section we want to show that the two theoretical tasks are very difficult: their accomplishment requires that we satisfy two desiderata that pull in opposite directions. Since the difficulties reveal themselves very clearly when you try to fulfil the task, this is what we will do: first we try to do it as good as possible (3.1 and 3.2); then we show that our attempt is not really satisfactory and point out that there is a tension between opposite desiderata (3.3).

3.1 Our attempt starts from a view on the aims of sciences which we call *sophisticated pragmatism.* The main idea is that the aim of science is not just to provide a true description of the world. In *Science, Truth and Democracy*, Philip Kitcher formulates an argument supporting this view:

> Nobody should be beguiled by the idea that the aim of inquiry is merely to discover truth, for, as numerous philosophers have recognized, there are vast numbers of true statements it would be utterly pointless to ascertain. The sciences are surely directed at finding significant truths. But what exactly are these? (2001: 65)

Kitcher's answer to the question at the end of the quote is double. On the one hand, there is *practical significance*:

> One possible answer makes significance explicitly relative – the significant truths for a person are just those the knowledge of which would increase the chance she would attain her practical goals. Or you could try to avoid relativization by focusing on truths that would be pertinent to anyone's projects – the significant truths are those the knowledge of which would increase anyone's chance of attaining practical goals. (2001: 65)

But for Kitcher there is more:

> Neither of these is at all plausible as a full account of scientific significance, and the deficiency isn't just a result of the fact that both are obviously rough and preliminary. Linking significance to practical projects ignores areas of inquiry in which the results have little bearing on everyday concerns, fields like cosmology and paleontology. Moreover, even truths that do facilitate practical projects often derive significance from a different quarter. Surely the principles of thermodynamics would be worth knowing whether or not they helped us to build pumps and engines (and thereby attain further goals). Besides the notion of practical significance, captured perhaps in a preliminary way by the rough definitions given above, we need a conception of "theoretical" or "epistemic" significance that will mark out those truths the knowledge of which is intrinsically valuable[5]. (2001: 65)

Because there is something more than just practical significance, we call a view on the aims of sciences like Kitcher's and ours *sophisticated pragmatism* as opposed to the view – which could be labeled *strict pragmatism* – that science should only look for practically significant truths (this view was defended by John Dewey).

Let us now apply this view to contrastive explanations. The general idea of sophisticated

[5] Note that the term "intrinsically valuable" is ill-chosen. What Kitcher means is: valuable for non-practical reasons.

pragmatism is that the aim of science is to find *significant truths*. Applied to the explaining of contrasts we get: the aim is to find *significant true* answers to why-questions about *significant true* contrasts. If we agree that this is the aim that scientists must have when they are explaining contrasts, this imposes restrictions on what scientists should do.

First, this aim implies a condition of significance for contrastive questions:

> It is worthwhile to attempt to answer the contrastive question "Why X rather than Y?" if and only if (a) X is true and (b*) the contrast is significant.

Let us compare this to (Q) from Section 2.2:

> It is worthwhile to attempt to answer the contrastive question "Why X rather than Y?" if and only if (a) X is true, (b) Y is false and (c) we have a relation R in mind that answers should and can satisfy.

The difference with our view is clear: our clause (b*) versus (b)+(c) in the erotetic model. As we have already mentioned in Section 2.2, we believe that the requirement that Y is false is clearly too strong and should be dropped. Moreover, while the theory extracted from van Fraassen's views is unable to avoid the green cheese problem, clause (b*) of our condition of significance of why-questions (which says that it is only worthwhile to answer a contrastive question if the contrast is significant) allows to cope with it. Of course, this does require us to say something about the notion of 'significance' (otherwise, we would only solve the problem by stipulation). We will do that in section 3.2.

The aim of finding significant true answers to why-questions about significant true contrasts also implies a restriction on answers:

> An answer to a contrastive why-question is an adequate explanation if and only if (a*) the question is about a significant true contrast, (b*) the answer is true and (c*) the answer is significant.

The corresponding claim (A) in Section 2.2 was:

> An answer to a contrastive why-question is an adequate explanation if and only if (a) the question is about a true contrast, (b) the answer is true and (c) it stands in the contextually determined relevance relation R to <X,Y>.

Again the difference concerns our referral to the concept of significance, both in (a*) an (c*). Where (A) provides no answer to the red herring problem, the extra clause of significance our proposal introduces does. This again raises the question: what is significance?

3.2 Our approach relies on the notion of significance to avoid excessive liberalism, both with respect to why-questions as well as answers. How to spell out this notion? There are two common ways in which a topic and foil may significantly relate.

First, it may be that the topic is contrasted with a foil that is *at odds* with the topic. These situations are common in science. To give an historical example, consider the question:

"Why does the orbit of Uranus deviate from the orbit predicted by the 1845 standard model of our solar system, while the orbits of all the other planets are as predicted by this model?" With "1845 standard model" we mean Newton's laws plus the assumption that there are 7 planets (Mercury till Uranus, but without Neptune). As is well known, Neptune was discovered in 1846 and its gravitational effect on Uranus explains the deviations. Borrowing some terminology from Nozick (1981), let us call these foils which make the topic rather unexpected *apparent excluders*. Apparent excluders do not operate in isolation: they make the topic unexpected if we combine them with (often implicit) background knowledge. For instance, in the Uranus case there is the assumption that there are no undiscovered planets.

Second, it may be that the foil represents a more *desirable situation* than the topic: it represents an improvement. Again, such situations occur frequently in science. For example, it is the case that the hole in the ozone layer over the Antarctic is expanding; since it not expanding is the more desirable situation, one might reasonably ask: "Why does the hole in the ozone layer over the Antarctic increase in size, rather than decrease or remain stable?" Such questions are asked with an eye to bring about a situation that is an improvement over the current situation.

The foregoing considerations give us enough material to construe a tentative[6] definition of significance with respect to why-questions:

> *Significance of contrastive questions*
> The contrast between topic X and foil Y is significant if and only if Y relates to X in one of the following ways:
> (a) Y is an apparent excluder of X, or
> (b) Y represents an improvement over X.

As we have seen in the previous subsection, Kitcher distinguishes between epistemic and practical significance. Option (b) is an implementation of the idea of practical significance for contrastive questions. Option (a) is an implementation of the idea of epistemic significance for contrastive questions.[7] A brief example: the question "Why is the level of unemployment in Belgium 7.4 %, rather than the moon made of green cheese?" is not significant according to our definition.

A similar tentative definition of significance can be developed for answers:

> *Significance of answers to contrastive questions*
> A statement A is a significant answer to a contrastive question Q if and only if
> (a) A dissolves an apparent exclusion in Q, or
> (b) A suggests a means of attaining a foil that is an improvement over the topic.

Again, the options (a) and (b) are implementations – for the case of answers to contrastive questions – of the ideas of epistemic and practical significance. A brief example: "Because

---

[6] This definition is tentative because we will point at its shortcomings in 3.3. However, we will not present a better one, since we don't have a better one; in that sense, the definition is not tentative: it is our best shot.

[7] There are other possible implementations of these ideas. We come back to this in Section 3.3.This is the most important reason why the theoretical tasks are so difficult.

red herrings have gills" is not a good answer to the (significant) question "Why is the level of unemployment in Belgium 7.4 %, rather than 10.7 %?" because it does not do any of the things specified in our definition of significance of answers.

Before we discuss their shortcomings, let us make clear that our two definitions certainly do not constitute an exhaustive theory of significance, for two reasons. First, we only deal with why-questions and explanations, while there are numerous other scientific activities for which one can try to define significance. Second, within the domain of why-questions and explanations, we have only considered contrastive questions: significance of non-contrastive why-questions is beyond the scope of this paper. In other words: the scope of our definitions is limited, so they cannot constitute an exhaustive theory of significance.8

3.3 Now that we have done our best to accomplish the theoretical tasks, let us see how well we did. An important characteristic of our definitions is that they are *disjunctive:* they tell us that there is significance if at least one of two conditions is fulfilled, but they do not tell us what the two possibilities have *in common*. In other words, they do not give insight into what significance *is*. They just list two possible cases. In logic and mathematics this is known as "definition by cases". Definition by cases may be a good instrument for representing our intuitions (if they lead to correct appraisals, see below) but they do not provide much philosophical insight. Even if the definitions do not suffer from counterexamples, they are not what we really want: because of their disjunctive nature, they do not tell us what the "nature" of explanation-seeking questions or the "nature" of explanation is.

A second problem with our definitions is that they are too restrictive: in order to avoid false negatives (intuitively significant questions and answers that are nevertheless excluded by our definitions) we have to add more disjuncts. To see this, let us adapt a question we used in 3.2. Suppose the hole in the ozone layer remains stable in year *x*. Then we can ask:

> Why does the hole in the ozone layer over the Antarctic remain stable this year rather than increase in size?

This question is similar to the one in 3.2, but topic and foil are switched. The underlying motivation is not improvement, because the foil is not the ideal state. A possible reason for asking this question is *preservation*: we want to preserve what is good.[9] This is a practical motivation for asking why-questions which is complementary to the one we have included in our definition (improvement). In order to avoid false negatives, we have to take such questions into account by adding a third possibility. The definition of significant answers must also be adapted.

The two problems together result in a tension. In order to have a definition which is philosophically interesting, we need to find something overarching so we can eliminate the

---

[8] It may also be useful to mention that our the definitions are meant to capture the meaning of significance. They should not be confused with an attempt to develop a "logic of significance" which would contain a formal test procedure for testing significance.

[9] Note that the topic (the hole in the ozone layer remaining stable) is also apparently excluded by our background knowledge of the hole expanding year after year. As such, this why-question may be asked with surprise in mind, besides preservation. But of course there is nothing against two interests overlapping: an apparently excluded state of affairs may also be a desirable one.

disjuncts. In order to account for our intuitions, we have to add more disjuncts. And more disjuncts make it more difficult to find an overarching principle that covers them all. Because of this tension, the theoretical task is extremely difficult. If someone accomplishes it, we will finally know what makes a contrastive question a real explanation-seeking one, and what makes an answer an explanation. However, because of the tension just explained, we are pessimistic about our chances.

4. Significant contrastive questions

As we saw at the end of Section 2, there are two practical tasks which philosophers of science can set themselves: formulate guidelines which help scientists to avoid uninteresting ("green-cheese-like") questions; and formulate guidelines which help scientists to avoid uninteresting ("red-herring-like") answers. In this section we develop and discuss guidelines of the first kind. The second type can be found in Section 5.

The guidelines for contrastive questions have the following form:

> If conditions c_1, …, c_n, hold then question X is a significant explanation-seeking question.

So the guidelines are positive: they are rules of thumb for formulating significant contrastive questions.

4.1 Here is the first guideline:

> Suppose that object *x* has property P at time *t.* Then the question
> "Why does *x* have property P, rather than the ideal property P*?"
> is a significant explanation-seeking question.

P and P* are mutually exclusive properties. We use the words 'ideal property' prescriptively, in the sense of it being (e.g. morally) better than the non-ideal property (as opposed to ideal in the sense of abstraction, like model idealizations). We call questions of the type considered in this guideline I-type questions because invoke an ideal state. Before we give an example and justify this guideline, let us present a second one which is closely related:

> Suppose that object *x* has property P at time *t* and object *y* (which like x belongs to class C) has the ideal property P* at *t.* Then the question
> "Why does *x* have property P, while *y* has the ideal property P*?"
> is a significant explanation-seeking question.

We call the questions referred to in this guideline I'-type questions because (like the I-type) they invoke an ideal state but in a different way.

Let us consider an example of each type. Suppose there are two fields of potatoes (*a* and *b*) which are regularly infected with late blight (*phytophtora infestans*, the oomcyete or microorganism that caused the 1845 potato famine in Ireland). Now about one of these fields, a researcher might pose an I-type question:

> Why has field *a* been struck by late blight, rather than remain healthy?

This type of question contrasts an observed state of affairs with an ideal one, i.e. one we consider to be preferable. According to our first guideline, it is an interesting explanation-seeking question. Suppose now that suddenly, during harvest time, it is found that the crops of field *b* have remained healthy, while field *a* is struck by late blight. In this

case, researchers might pose an I'-type question:

> Why has field *a* been struck by late blight, while field *b* remained healthy?

This question contrasts two observed state of affairs, one of which is the ideal one. The two objects have something in common: they both are crop fields. According to our second guideline, this is a significant explanation-seeking question.

Let us now turn to the justification of the guidelines. The two guidelines we have presented here follow from the fact that *improvement* is an *epistemic interest.* Epistemic interests are types of motivations for (a) scientists to search for explanations and (b) other people (policy makers, the general public) to be interested in the explanations scientists give and to pay them for their research. Improvement in the sense of making a given situation better is certainly one of these typical motivations. *Prevention* (which is closely related to improvement, but more future-oriented) and preservation (see 3.3) are also epistemic interests. A fourth one is *attributing moral and/or legal responsibility*: this is often the underlying epistemic interest when we explain actions (see Weber & Vanderbeeken 2005 for this).

4.2 Here is a third guideline:

> Suppose that object *x* has property P at time *t* and objects $y_1, \ldots y_n$ (which like *x* belong to class C or bear relation R to *x)*[10] have property P*. Then the question
> "Why does *x* have property P, while objects $y_1,\ldots, y_n$ have property P*?"
> is a significant explanation-seeking question.

We call questions of the format used here "E-type" questions. The fourth guideline is:

> Suppose that object *x* has property P at time *t* and property P* at time *t'*, and we have a reason to expect no evolution in *x* in this period. Then the question
> "Why does *x* have property P* at *t'*, while it had property P at *t*?"
> is a significant explanation-seeking question.

We call questions of the format used here T-type questions. E- and T-type questions have foils that apparently exclude or diminish the chance of the topic occurring.

---

[10] Most question-types we consider in this article contrast properties of the same object. The E-type question however, compares two objects in their not sharing certain properties. Nevertheless, these objects are not chosen randomly, for this would mean inviting back in nonsensical questions again. The extra clause 'which like x belong to class C or bear relation R to x' is used to indicate that the objects stand in some kind of relation to each other, or share membership of a certain class. Nevertheless, if there are no constraints on C, this would again make our position vulnerable to green cheese-type questions. While we do not believe there are any principal constraints on C, we do believe that there are context-sensitive constraints – in particular, that the objects compared are similar in a sense which is relevant to the explanation-seeking question, as we will momentarily show when we illustrate the E-type question with an example. The ways in which context determines when two or more objects are similar enough to motivate E-type questions is deserving of further research.

Let us now give an example. Suppose it is observed that a western European country *c* has no increase in unemployment while its neighbours have. Then we might pose an E-type question:

> Why did the unemployment figure of *c* remain stable given that it rose in the surrounding countries?

Given the increase in the surrounding countries, we had expected the figure to rise in *c* as well. That is, the foil is an apparent excluder of the topic, and hence, the question is significant. The answer to this question then should tell us where our expectations went wrong.

Suppose that it is further observed that this rise in unemployment in the surrounding countries has happened quite suddenly. We might then ask a T-type question (*d* is one of the neighbours):

> Why has the unemployment figure of *d* risen during the last year, while it remained unaltered the year before?

This question contrasts a current state of affairs with a past one that is not the case anymore. In this example it is an unexpected evolution in time that is the target for explanation.

Note that one could also ask a question about the *non*-occurrence of a certain evolution. In the example under consideration, if the unemployment figure has not risen, while we had reasons to believe it would (for example if there was a monetary crisis going on), then it is the *absence* of change which is in need of explanation. Let us call that a T'-type questions. In both cases (T and T') it is surprise *given some preconceived opinions or arguments* that prompts the question. The transition between the two periods can be marked by change or continuity, but is always unexpected. Similarly, we can define E'-type questions as the complement of E-type questions: E'-type questions are about unexpected similarities between objects or events.

The justification of these guidelines again relies on epistemic interests. In 4.1 the epistemic interests were of a practical nature. Here we have to invoke more theoretical ones, such *resolving apparent exclusions* or more generally, the desire to unify our knowledge and make it more coherent (resolving apparent excluders is a specific way to increase coherence).

4.3 We certainly do not claim that the guidelines we present are jointly exhaustive. One of the reasons why we do not claim exhaustiveness is that we focus on contrastive questions about particular facts. Similar guidelines can be formulated for other types of why-questions. And even if we restrict ourselves to contrastive why-questions, many more positive guidelines could be developed.

Before we turn to significant answers, let us briefly come back to the difference between the theoretical and practical tasks. The theoretical tasks requires filling in a biconditional, the practical tasks are about sufficient conditions. Since the latter are easier to find, it is possible

to speak of "grand" task instead of theoretical and "modest" instead of practical.[11] There are two reasons why we prefer our terminology:
(1) The task belong to a different type of philosophy of science; the practical task belongs to "philosophy of scientific practice", which tries to be useful for scientists; the theoretical task belongs to a more reflective philosophy of science which looks back at what scientists are doing and reflects upon it.
(2) As emphasised in Section 3.3, definition by cases is not adequate in order to fulfil the theoretical task. There is more to the theoretical task than the biconditional. This qualitative difference is better expressed by our terminology.

5. Significant answers

Let is now turn to guidelines for answers. We will show that it is possible to formulate guidelines of the following form:

> If the question is a significant contrastive explanation-seeking question, and the motivation for raising it is of type X, then an answer that satisfies condition Y is a significant answer.

In such guidelines "condition Y" is the relevance criterion that is supposed to ensure that the scientist who follows the guideline gives a significant answer. As we will see, Y can be complex (made up of several subconditions). Note that our format refers types of motivations for asking why-questions (epistemic interests). Their role will become clear in what follows.

Many guidelines of this form can be developed. In Section 5.1 we will propose a guideline for I- and I'-type questions. In Section 5.2 we present one for T-type questions.

5.1 According to an article in Time Magazine of January 1, 1945, the US army, fighting on the European continent in the last wet months of the previous year, suffered an outbreak of *foot immersion syndrome*, also known as trench foot (after its devastating impact at the western front during the First World War). Trench foot is a condition brought on by poor vascular supply to the feet, due to prolonged exposure to damp and cold conditions. Symptoms include numbness, swelling and early stages of necrosis in affected areas. According to the article, approximately 17,500 U.S. soldiers developed the condition. An I-type question would be:

> Why did the American soldiers suffer from trench foot, rather than stay healthy?

The epistemic interest which motivates this question might be *improvement*: we might be interested in keeping the soldiers healthy, and discovering the cause of trench foot is of

[11] We thank on of the referees for pointing at this possibility.

course instrumental in doing so.[12] This last point is important: in posing this type of question with a therapeutic motivation, the researcher determines that – in order to be significant – an answer to his particular question must have *causal relevance* to the contrast in that question. In the case at hand, causal factors in developing trench foot are prolonged exposure of feet to cold and damp conditions; conditions that are typical of northwest Europe during November and December.

But there is more. The article goes on to say that in contrast, the British army, even though it made its way through the damp plains of Holland, reported no similar problems, as its soldiers wore robust gum boots, which they were required to keep waxed, and were instructed to regularly massage their feet and change their socks. In general, the British were ordered to take care of their feet, while the Americans were not. This information makes it possible to pose an I'-type question:

> Why did the American soldiers suffer from trench foot, while the British soldiers did not?[13]

The motivation behind this I'-question can be improvement, but the added bonus over the I-type question is that the answer you get to the former typically suggests that there is a solution which is not beyond our reach.

The explanation might look like this:

> The American army experienced an outbreak of trench foot because:
> 1) the conditions were damp and cold; and
> 2) the soldiers did not take care of their feet.
>
> The British army did not experience an outbreak of trench foot even though:
> 1) the conditions were damp and cold, because
> 2) the soldiers took care of their feet.

We will use this example to illustrate our first guideline for answers:

> If an I-type or I'-type question is asked, and the motivation for asking it is "improvement", then an answer that (a) gives causal information, (b) highlights causes that make a difference with respect to the contrast, and (c) at least partially uses causes that relate to possible human interventions, is a significant answer.

This guideline contains three conditions. Let us look their justification. That the answer has to give causal information follows immediately from "improvement" interest which is in the

---

12 Of course, this is not to say that the journalist who actually wrote the article had any interest beyond simply reporting an outbreak of trench foot (although, being an American himself, he would in all likelihood not have been entirely neutral). However, Army officials may very well have this interest.

13 Here we see an example of a significant question that is precluded by van Fraassen's demand that the foil is false. Hence our remark in section 2.2 that we think this demand is too strict. In some cases, scientists compare two states of affairs which are both true.

antecedent part of the rule. Condition (b) is necessary because not all causal information is relevant. Peter Lipton gave a nice example to show this:

> Suppose that my car is belching thick, black smoke. Wishing to correct the situation, I naturally ask why it is happening. Now imagine that God (or perhaps an evil genius) presents me with a full Deductive-Nomological explanation of the smoke. This may not be much help. The problem is that many of the causes of the smoke are also causes of the car's normal operation. Were I to eliminate one of these, I might only succeed in making the engine inoperable. (1993, p.53)

The example which Lipton considers here is an I-type question. His message is clear: the answer should focus on causes that make a difference; it makes no sense to present and describe all the causes.

Finally, the condition that at least one of the causes must relate to a possible human intervention (like the soldiers taking or not taking care of their feet) directly follows from the "improvement" motivation: if there is no such cause, we cannot ameliorate the situation.

5.2 Next, we will consider significant answers to T-type questions. Recall that the foil is, by definition, always an apparent excluder of the topic. So there is some surprise and unexpectedness involved.

We will discuss answers to T-type questions with an example of the social sciences: the explanation of the *gender gap.* The term *gender gap* refers to differences between men and women on various public and private issues in the political and social sphere. Specifically, social scientists have been interested in how these gender related differences play out in the respective voting behaviour of men and women. During the 1950s and 1960s, men and women in the US exhibited very similar voting behaviour, as both tended to favour right-wing parties (there was no significant gender gap); however, from the 1980s on, this pattern changes as women were reported to offer disproportionate support to left-wing parties compared to men (Chaney, Alvarez & Nagler, 1998; Manza & Brooks 1998).[14] Thus, the gender gap occurred, constituted by a shift in voting behaviour of American women over time as compared to men. About this phenomenon, researchers asked a T-type why-question of the form:

> Why did American women offer more support to left-wing parties in the 1980s and 1990s, while on average they voted more conservative in the 1950s and 1960s?

Why did researchers ask this question? It was motivated by surprise: they had reasons to believe that no change would occur. In fact, early post-war sociology did not focus on gender at all (Stouffer 1955; Lipset 1960), the general view among social scientists being that

---

14 Thus, the gender gap refers specifically to a gender based divergence in voting behaviour, not to the way this divergence is actually situated on the political spectrum: if women had continued to vote conservative while men changed to liberal, this would also have constituted a gender gap. Of course, the explanation would be quite different in such a case.

women simply lacked interest in politics (Berelson et al. 1954, p. 25). It was thought that when women did cast their vote, their acceptance of the traditional family roles simply made them follow their husband's choice (Campbell et al. 1960). Given these preconceived opinions, the T-type question they asked was motivated by surprise: an unexpected evolution in time had to be explained. Therefore, a significant answer to a question like this should tell us why things are different from how we expected them to be.

The answers researchers provided cited causes that were located at the level of U.S. politics. That is, social scientists tried to explain this shift in political preferences by referring to factors specific to the U.S. political situation, such as party differences on the Equal Rights Amendment (a proposed amendment to the American constitution guaranteeing equal rights for women under federal, state and local law) and strong divisions on topics as abortion and welfare reform policies (Costain & Berggren 1998; Mueller 1988).

This example illustrates the following guideline:

> If a T-type question is asked, then an answer that (a) gives causal information, (b) highlights causes that are present in the first period and absent in the second (or vice versa) is a significant answer.

The justification for this guideline is that the two conditions ensures that the apparent exclusion in the question is dissolved.

5.3 It is worthwhile to explore this example about the gender gap a bit further. The explanations scientists gave for gender gap in the U.S. leads to further expectations: insofar as the causes of the gender gap are specific to the political situation in the U.S., we would not expect to find a similar pattern in other countries, where these factors are absent.

However, Giger argued that evidence obtained from the EuroBarometer suggests that in most Western-European countries, a similar shift in political preference among women is discernible, though appearing somewhat later and in varying degrees across countries (Giger, 2009). Thus, Giger was led to ask a T-type question:

> Why does voting behaviour among women of Western European countries change, rather than remain unaltered?

This question was also motivated by surprise: given the fact that the gender gap was previously explained by referring to conditions specific of the U.S., the result was unexpected for everyone who accepts that explanation. A significant answer to this question should ideally indicate what led us to the wrong expectation: it should refer to factors that were previously ignored. Giger argues that the developmental theory of gender realignment (henceforth DTGR) by Inglehart and Norris (2000, 2003) does a better job of explaining the phenomenon of the gender gap, to the extent that it also explains its occurrence in Western Europe. Rather than pointing at specific circumstances of any one country, DTGR points at structural and cultural developments that are common to wealthy, post-industrial societies. These include: reforms in the paid labour force for women, more equal opportunities of education and a shift in traditional family values. By referring to these factors, the theory explains why the gender gap also manifests itself in Western European countries and the

developed world in general (rather than post-communist or third world societies), and reveals what features were left out the original explanation and why this omission led us to wrong expectations.[15]

6. Hands-on contextualism

6.1 Khalifa (2004) makes an observation we also made in this article, namely that van Fraassen's theory of explanation is too thin to exclude the kind of ridiculous questions and answers exemplified by the green cheese and red herring problems. Let us first check what his aim is and how it relates to the theoretical and practical tasks we have distinguished. According to Khalifa, the challenge that results from the green cheese problem is "to offer principles for narrowing down the contrast class" (2004, p. 45). This sounds like a the first practical task we have distinguished. Khalifa's response to the challenge consists in three *strategies* (his term) which can be used to arrive at good questions (see 6.2). This is clearly an attempt to fulfil the practical task we have described. Khalifa's paper does not contain a definition (comparable to what we did in Section 3) of significant questions. So he does not attempt to achieve the first theoretical task. That is also the case for the answers to the questions: Khalifa gives guidelines, his paper does not contain a definition of what an explanation is. So he does not want to fulfil the second theoretical task either.

6.2 Khalifa presents three strategies for identifying foils. First, the topic and foil tend to refer to the same thing; this may be called the *semantic strategy*. Second, the topic and foil are confirmable by the same data-generating procedure; let us call this the *methodological strategy*. Third, the most relevant foils will be those that are the expected results or implications of a plausible hypothesis; we might label this the *epistemic strategy*. For Khalifa, these three strategies are complementary: together they identify relevant foils, and it is their complementary application that allows scientists to avoid "green-cheese-like" questions.

Let us now compare the strategies to what we have done in Section 4. We begin with the epistemic strategy: the idea that the most relevant foils are the expected results or implications of a plausible hypothesis. How does this work? ? In his discussion of Blau and Duncan's *The American Occupational Structure* (1967), Khalifa writes:

> Blau and Duncan provisionally began with the hypothesis that African American social mobility was analogous to white social mobility but revised it as the data bore out certain differences. Whites coming from the lowest socioeconomic origins had the highest chance of upward mobility. From this Blau and Duncan conjectured that African Americans should, on average, be more likely to experience upward social mobility than whites since a larger portion of the

15 Which is not to say that the U.S. specific explanation contradicts the one provided by Giger: for all we know, the U.S. specific factors are still explanatorily relevant for explaining the gender gap in the U.S. In that sense, the first answer, though it led us to entertain wrong expectations, was significant in its own right. Rather, Giger's explanation is better in that it highlights different causal factors that are *also true* of the U.S. That is to say, these causal factors might have been used to explain the gender gap in the U.S., and if they had been, we would not have been led to entertain false expectations (namely that the gender gap is specific to the U.S.).

> African American community was concentrated in these lower socioeconomic strata. However, the data confirmed quite the opposite […] From this, we get the following why-question: Why do African Americans have a lesser (rather than greater) chance of upward mobility than white Americans?" (2004, p. 46).

One can easily see the similarities between this situation and the one we described in the explanation of the gender gap: given a certain situation, we come to harbour certain expectations, and when these expectations are disconfirmed, this leads us to pose a contrastive why-question that is motivated by surprise. As Khalifa says: "…expectations play a crucial role in identifying which propositions will be included in the contrast class" (2004, p. 47). Obviously, this epistemic strategy for choosing foils constitutes a similarity between our proposal and Khalifa's hands-on contextualism: our third and fourth guideline in Section 4 correspond to this idea.

However, there are two important differences. The first is that we have templates for our guidelines and that we use question-types to formulate specific guidelines. So we have a *general, unified* strategy for avoiding non-significant questions: guidelines with certain conditions in the antecedent and questions of certain types as output. Khalifa rightly stresses that we have to avoid nonsensical why-questions such as "Why is George W. Bush president rather than George W. Bush president?' in which topic and foil are identical (2004, p.46). Our types guarantee this. Khalifa's semantic and methodological strategy are meant to avoid too much variation. Our question-types avoid excessive variation between topic and foil in two ways:

(1) the objects are identical or similar (as they were in the examples of the crop fields and the American and British armies)

(2) P and P* are mutually exclusive, which entails that they belong to the same family of predicates.

In this way, our question types can incorporate the semantic and methodological strategy superfluous: they enable us to avoid the green cheese problem without using these strategies. The underlying idea of the semantic strategy (shared reference) is incorporated in (1), while the fact that P and P* belong to the same family of predicates implies that they can be measured by the same methods (so the methodological strategy is also incorporated in an indirect way).

The second additional difference is that, while Khalifa only considers expectations in his epistemic strategy, our approach also includes two guidelines based on ideals. As we saw in the examples of I-type and I'-type questions, sometimes the relevant contrast is between a state of affairs which we know to be true and one which we *desire* to be true; i.e. between a factual and an ideal situation, rather than an unexpected one. Our theory is more complete because it adds this important element (Khalifa could have included this idea in his epistemic strategy, but he does not do that).

6.3 With respect to the practical red herring task (helping scientists to avoid uninteresting answers), Khalifa clarifies how relevance criteria are constrained by the reliability and soundness of methodological conventions. In the case of Blau and Duncan's 1967 study of occupational status, the relevance criteria are narrowed by the data-generating procedures – procedures providing a normative basis for the features an acceptable answer/explanation

must have. These procedures involved the measurement and statistical analysis of occupational status and of social mobility. Blau and Duncan generated data on occupational status by defining it as a function of the median income and median education (and social mobility data were parasitic on the occupational status data). They considered their account an advance in comparison with its predecessors like the one's based on the data of the *National Opinion Research Center* (that traced the attitudes people harbor toward occupations), because it "showed why income and education were the *causes* of occupational status", according to Khalifa, and an "explanation that can identify a plausible causal mechanism will, ceteris paribus, be more relevant than one that does not". (2004, p. 50)

We agree that the identification of causes, mechanistic details and isolable variables could be helpful in providing significant answers to, for instance, I- or I'-type questions. In that sense, Khalifa's discussion provides an example that is in line with our account. However, our guidelines are both more general, in that they are not merely after causal mechanisms, and more specific, because they specify exactly what kind of causal information a significant answer should contain. This is possible because of the use of question-types and epistemic interests.

7. Conclusion

In this paper we have shown that the green cheese problem and the red herring problem each lead to two tasks that philosophers of science can set for themselves: a theoretical task and a practical task. We have argued that the two theoretical tasks are very difficult: their accomplishment requires that we satisfy two desiderata that pull in opposite directions.

In order to make some substantial contributions to the two practical tasks we have formulated and defended guidelines for explanation-seeking questions and for answers. We have shown that these guidelines incorporate the core ideas of Khalifa, but have important advantages.

References

Berelson, B. R., Lazarsfeld, P. F. & McPhee, W. (1954). *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.

Campbell, A., Converse, P. E., Miller, W. & Stokes, D. E. (1960). *The American Voter*. New York: Wiley.

Chaney, C. K., Alvarez, M. R. & Nagler, J. (1998). Explaining the gender gap in U.S. presidential elections, 1980-1992. *Political Research Quarterly* **51**, 229-311.

Costain, A. N. & Berggren, H. (1998). The gendered electorate. Paper presented at the annual meeting of the American Political Science Association, Boston, September.

Giger, N. (2009). Towards a modern gender gap in Europe? A comparative analysis of voting behaviour in 12 countries*. The Social Science Journal* **46**, 474-492.

Inglehart, R. & Norris, P. (2000). The developmental theory of the gender gap: Women's and men's voting behaviour in global perspective. *International Political Science Review/Revue internationale de science politique* **21** (4): 441-463.

Inglehart, R. & Norris, P. (2003). *Rising tide. Gender equality and cultural change.* Oxford: Cambridge University Press.

Khalifa, K. (2004). Erotetic contextualism, data-generating procedures and sociological explanations of social mobility. *Philosophy of the Social Sciences* **34**: 38-54.

Kitcher P. & Salmon, W. (1987). Van Fraassen on explanation. *Journal of Philosophy* **84** (6): 315-30.

Kitcher, P. (2001) *Science, Truth and Democracy.* New York: Oxford University Press.

Lipset, S. M. (1960). *Political Man*. Baltimore: John Hopkins University Press.

Lipton Peter (1993), 'Making a Difference', in *Philosophica* **51**, pp. 39-54.

Manza, J. & Brooks, J. (1998). The gender gap in U.S. presidential elections: When? Why? Implications*? American Journal of Sociology* **103** (5): 1235-1266.

Mueller, C. (ed.) (1988). *The Politics of the Gender Gap.* London: Sage.

Nozick, R. (1981). *Philosophical Explanations*. Cambridge MA: Harvard University Press.

Risjord, M. (2000). *Woodcutters and witchcraft: Rationality and interpretive change in the*

social sciences. Albany, New York: State University of New York Press.

Stouffer, S. (1955). *Communism, Conformity, and Civil Liberties*. New York: Doubleday.

Time magazine U.S. (1945). Medicine: Again, Trench Foot. Monday Jan. 1. Available online at: http://www.time.com/time/magazine/article/0,9171,791832,00.html

van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.

Weber, E. & Vanderbeeken, R. (2005). The functions of intentional explanations of actions. *Behavior and Philosophy* **33**: 1-16.