The surprise exam

PHIL 20229
Jeff Speaks

April 8, 2008

1 An argument for the impossibility of surprise exams

Suppose that I come into class and announce the following revision to the schedule: in addition to the final paper and final exam, we will have one pop quiz for 99% of your grade on some class day between now and the end of the semester. (The topic will also be a surprise.) I won't tell you which day I am going to give the exam, but I will tell you this: I will definitely give an exam on one of the remaining class days, and on that day you will have no good reason to believe that it will be on that day, rather than some other. This is just what it means for it be a surprise exam, of course.

Having taken a course in Paradoxes, you immediately realize that I have said that I am going to do something impossible, and reply to me as follows:

> Well, you can't give the exam on the last day of class, since then we would know that, there being no class days remaining, you had to give the exam that day; and in that case it would not be a surprise. So we can safely eliminate the last day of class from the list of possible days on which you can give the exam.

> But then you can't give it the second to last day of class, either; for on that day we would know that you couldn't wait till the last day — since then it would not be a surprise —- and so we would know that you had to give it that day. But then, of course, it would not be a surprise. So we can also safely eliminate the second-to-last day of the semester as a possible date for the exam.

> But then you can give it the third to last day . . .

> (and so on.)

Something is wrong with your line of reasoning, since I can clearly give you a surprise exam — but what is it?

2   The announcement must be known by the class

We need to be a bit clearer about the nature of my announcement to the class, and about what exactly is supposed to be paradoxical about it.

A first try is to say that my announcement to the class seems as though it should be true; but the reasoning offered above seems to show that it is is false. But this is not quite right; there is no problem with an announcement of the sort described above being true; if the class does not believe the announcement, then they will certainly be surprised by the exam! So we need to suppose that the announcement is both true and known to be true by the class — on this view, the paradox is that the line of reasoning pursued above shows that the announcement cannot be both true and known by the class to be true on each day of the semester.

Is this still genuinely paradoxical? Could we simply rest by saying that the teacher can give a surprise exam, but only if either the class does not know that he will, or the class fails to see the argument sketched above? We aren't, after all, saying that the teacher has said something which cannot be true.

Does it make sense to say that something can be true but unknowable? Is that required to adopt the present response to the paradox?

A separate response to this version of the paradox focuses on a questionable step in the reasoning sketched above. We are now imagining that the class knows that

> The teacher's announcement is true.

We imagined that the class reasons to the conclusion, which they then know, that

> If we know that the teacher's announcement is true, then the exam cannot be on the last day of the semester.

They are then supposed to combine these two pieces of knowledge to give them the knowledge that

> The exam cannot be on the last day of the semester.

But how is this step supposed to work? Ordinarily, we think that one can move from the knowledge that p and the knowledge that if p, then q to the knowledge that q. But the above reasoning is not of that form; instead we are moving from the knowledge that p and the knowledge that if <u>I know that</u> p, then q to the knowledge that q. This seems to make sense only if we assume not only that the class knows that the teacher's announcement is

true, but also that they known that they know that the teacher's announcement is true. Can we assume, in general, that if we know that something is the case, we can also know that we know that it is the case?

3 A self-referential version of the paradox

Above we noted that one might respond to the above version of the paradox by saying that the teacher's announcement cannot be both true and known by the class; nothing in the case shows that the teacher's announcement can't be true. But we can change the announcement in a way which does seem to give us this result. Imagine that the teacher's announcement to the class is the following:

> The self-referential announcement
>
> There will be a pop quiz some day between now and the end of semester, you know that this announcement is true, but you cannot know on the basis of this announcement when this pop quiz will occur.

If the reasoning sketched above is correct, then it seems that we have an argument that the self-referential announcement is not possibly true: it is a concealed contradiction, like 'there is a town with a barber who shaves all and only those townspeople who do not shave themselves.'

The self-referential announcement, as its name suggests, has a curious feature: it is an announcement which refers to itself, like 'This sentence is the one that I am uttering right now.' These sorts of self-referential utterances can be used to formulate a related paradox about knowledge.

4 The knower

Suppose that the teacher says to the class

> You, class, know that this very announcement is false.

Let's call this sentence THE KNOWER. Then we can give the following argument:

1. The class knows that this announcement is false.
2. This announcement is false.
3. _____
C.