

## SURPRISING THE EXAMINER

Michael WREEN

Over thirty years have passed since the so-called Paradox of the Surprise Examination<sup>(1)</sup> first saw print;<sup>(2)</sup> but although there have been numerous comments on and/or attempts to solve, dissolve, or regain the paradox,<sup>(3)</sup> I have yet to see a discussion I find fully

(<sup>1</sup>) Actually, the same "paradox" can be conveyed by any of a number of anecdotes, some of which concern a surprise examination, some a surprise air-raid drill, some a surprise execution, some an unexpected egg, etc. In what follows I shall concentrate on the "paradox" in its most discussed form, as involving a surprise examination.

(<sup>2</sup>) D. J. O'CONNOR, "Pragmatic Paradoxes," *Mind* (1948): 358-359.

(<sup>3</sup>) L. J. COHEN, "Mr. O'Connor's Pragmatic Paradoxes," *Mind* (1950): 85-87; Peter ALEXANDER, "Pragmatic Paradoxes," *Mind* (1950): 536-538; Michael SRIVEN, "Paradoxical Announcements," *Mind* (1951): 403-407; D. J. O'CONNOR, "Pragmatic Paradoxes and Fugitive Propositions," *Mind* (1951): 536-538; Paul WEISS, "The Prediction Paradox," *Mind* (1952): 265-269; W. V. QUINE, "On A So-Called Paradox," *Mind* (1953): 65-67; reprinted under the title "On A Supposed Antinomy," in W. V. QUINE, *The Ways of Paradox and Other Essays* (Cambridge, Mass., 1976), pp. 19-21; R. SHAW, "The Paradox of the Unexpected Examination," *Mind* (1958): 382-384; Ardon LYON, "The Prediction Paradox," *Mind* (1959): 510-517; David KAPLAN and Richard MONTAGUE, "A Paradox Regained," *Notre Dame Journal of Formal Logic* (1960): 79-90; G. C. NERLICH, "Unexpected Examinations and Unprovable Statements," *Mind* (1961): 503-513; Brian MEDLIN, "The Unexpected Examination," *American Philosophical Quarterly* (1964): 66-72; Frederic FITCH, "A Godelized Formulation of the Prediction Paradox," *American Philosophical Quarterly* (1964): 161-164; R. A. SHARPE, "The Unexpected Examination," *Mind* (1965): 255; J. M. CHAPMAN and R. J. BUTLER, "On Quine's 'So-Called Paradox'," *Mind* (1965): 424-425; James KIEFER and James ELLISON, "The Prediction Paradox Again," *Mind* (1965): 426-427; Jonathan BENNETT and James CARGILE, assorted reviews, *The Journal of Symbolic Logic* (1965): 101-103; Judith SCHOENBERG, "A Note on the Logical Fallacy in the Paradox of the Unexpected Examination," *Mind* (1966): 125-127; J. A. WRIGHT, "The Surprise Exam: Prediction on the Last Day Uncertain," *Mind* (1967): 115-117; James CARGILE, "The Surprise Test Paradox," *The Journal of Philosophy* (1967): 550-563; Robert BINKLEY,

satisfactory. In this paper I shall first present the case of the surprise examination in a slightly more elaborate form than has hitherto appeared in the literature. Having done so, I shall go on to discuss and to analyze critically all of its important components.

## I

On the first day of class a particularly tough instructor of philosophy told his Advanced Logic students that

If I say anything false this semester,  
and you, thinking on your own, detect  
my error and bring it to my attention  
in private, then you'll be awarded an  
"A" on the next exam.

The teacher, however, was an especially intelligent and careful man, and he asserted nothing false during the first half of the semester.

Then, just after mid-semester recess, he made the following cryptic remark:

There will be a surprise examination  
sometime next week.

When asked to explain what he meant by "surprise," he replied that he meant "unexpected" in the strongest sense of the term. He meant, in other words, that

"The Surprise Examination in Modal Logic," *The Journal of Philosophy* (1968): 127-136; A. K. AUSTIN, "On the Unexpected Examination," *Mind* (1969): 137; J. McLELAND, "Epistemic Logic and the Paradox of the Surprise Examination," *International Logic Review* (1971): 69-85; A. J. AYER, "On A Supposed Antinomy," *Mind* (1973): 125-126; B. H. SLATER, "The Examiner Examined," *Analysis* (1974): 49-50; Crispin WRIGHT and Aidan SUDBURY, "The Paradox of the Unexpected Examination," *Australasian Journal of Philosophy* (1977): 41-58; Igal KVART, "The Paradox of Surprise Examination," *Logique et Analyse* (1978): 337-344; Peter GALLE, "Another Note on the 'Surprise Test' Puzzle," *Informal Logic Newsletter* (1981): 21-22.

The class would be tested some day next week, and you, speaking distributively now, will have absolutely *no* expectation, either now or just prior to the exam, that the exam will be given that day.

His students, sharp logicians all, thought that they had guaranteed "A"s on the next exam; for, with one noteworthy exception, a student named McQ, all thought that the teacher had contradicted himself, and so spoken falsely.

"What could be more absurd!" all but the one thought. "A school-week is five days, Monday through Friday, so we have only five class meetings to worry about. But he can't give us the exam on Friday, because if he waited that long, there would be only a day left, Friday, so we'd all know it was coming, and it wouldn't be unexpected. Friday is out, then. But then again, he can't give it on Thursday either, can he? With Friday eliminated, the last day the test could be is Thursday; but since Thursday *is* the last day, if it were given on that day, then once again I would know it was coming, the exam would be expected, and it wouldn't be a surprise at all. So Thursday is out, too, and Wednesday is the last possible day. But not even that's true, now that I think about it. The reasoning I've used to eliminate Friday and Thursday eliminates Wednesday as well, doesn't it? – and then, as far as I can see, also eliminates Tuesday and finally Monday as possible exam days. The teacher, therefore, *can't* give us a surprise exam next week – which means that he's contradicted himself, which means that he's spoken falsely, which means that I'm guaranteed an 'A' on the next exam, because I'm sure going to hold him to that claim he made the first day of class."

Every student except McQ being very confident of the above line of reasoning, each said nothing at the time but waited for his chance to speak to the instructor in private. Unfortunately, the teacher had made sure that he would be very busy for the next week, and was unavailable for even the briefest of private meetings – though not for class – during the next week.

On the Wednesday following his announcement, the instructor strode into class, called roll, handed out blue books, wrote a number of questions on the board, and said:

Here is your surprise examination.

His students, again all of them save McQ, protested and tried to show him that he couldn't do such a thing, since a surprise exam was impossible. Without waiting for the students to complete their argument, the teacher snapped:

I said that you would have a surprise exam within the next week. The exam is on the board, it is being administered within a week of my announcement, and all of you – with one exception – are surprised. A surprise examination is not an impossibility – it's an actuality.

None of the students had prepared for the exam, so all flunked. All except one. McQ. He got an "A".

## II

The two central questions which need to be asked of the case are: (1) Where and how did the students go wrong, if in fact they did go wrong? and (2) How did McQ manage to get an "A" on an exam he hadn't prepared for? One way to find out the answers is to continue the story.

## III

Shortly after the exam had begun, the teacher asked McQ to speak with him in his office. McQ agreed, and the following is a report of their conference.

"Well, McQ," the instructor began, "you get an 'A' on the exam, but just for my own peace of mind, and in order to be sure that you really deserve the grade, I'd like you to explain your reasoning." Being anything but reticent, McQ was happy to oblige him.

"Well, the first thing I did was to reason along what must have been

the same lines that the other students did," said McQ, who then recounted exactly their line of reasoning. "But the argument seemed too pat, too easy," he continued, "and since you hadn't said anything false up to that point in the semester, I suspected that there was something wrong with it. Since it couldn't do any harm to explore my suspicion, I decided to think a little bit more about the exam.

"The first thing I did was to take a look at the conclusion of the argument, that a surprise exam next week is impossible. Remembering your exact words, I glossed that conclusion as: it's impossible (1) that there will be an exam next week, and (2) that the exam will be not be expected, will be *totally* unexpected, by me, considered as an individual student. 'Could (1) and (2) both be true?' I asked myself. And the answer was immediate: of course they could. If, for instance, you hadn't spoken at all, I thought to myself, then had you given an exam on Friday, both (1) and (2) would have been true – and in fact you *did* give an exam today, Wednesday, which was a total surprise to everyone except me. Anyway, the thought experiment I conducted convinced me of two things. First, the conclusion of the test-elimination argument is false, so there must be something wrong with it. Second, since in my thought experiment you said nothing, I was led to think that the error might have something to do with your having actually said something to the class last week. How exactly the error was connected with that fact I didn't know right then – but, as you'll see, I did after I had examined the argument further.

"Taking the conclusion as false, then, I began to take a closer look at the test-elimination argument itself. As far as I could see, the argument has two main stages. In the first, there is an argument leading up to the elimination of Friday. In the second, it is claimed that *if* the reasoning which eliminates Friday is sound, then the same reasoning eliminates every other 'last day.' The antecedent of this conditional having already been affirmed, Thursday through Monday are eliminated. Well, there didn't seem to be anything wrong with the conditional, so I thought the trouble must be with the first stage of the argument, with the sub-argument which eliminates Friday.

"I knew that if the trouble was there, that means that the same problem could have arisen if you had said, 'There will be a surprise exam next Friday.' After conducting another thought experiment, I

convinced myself that that was indeed so.<sup>(4)</sup>

"But what then is wrong with the argument which eliminates Friday? The answer, it seemed to me, is that one of the premises (explicitly) used in that argument, namely

(K) If the exam is administered  
next Friday, then I know, or at  
least I will know before Friday,  
that it will be administered then

is false. But that left me with two questions: (a) Why did the other students think (K) true? and (b) Why is it false?

"The first thing that occurred to me in relation to (K) was that with most cases of a proposition's being true, its being true is not a sufficient condition for my knowing it now or ever. Most substitution instances of the principle, If p, then I know or will know that p, are simply false. And I think that the other students knew as much too. So why were they seduced by (K)?

"The answer goes back to something I mentioned earlier: you said what you did where and when you did – you assertively uttered 'There will be a surprise exam next week' in the presence of a number of intelligent and prudent students. When a person assertively utters a sentence in the presence of others, and when that person is (or at least is thought by his listeners to be) in a position to know the truth of the proposition asserted, and when that person is (or at least is thought by his listeners to be) generally sincere, honest, and reliable in the kind of circumstances in which he does utter that sentence, and when his listeners are not otherwise in a position to know the truth of the proposition asserted, etc. – then his listeners will (normally) believe that the proposition in question is true. The students, for that reason, believed that there would be an exam. Knowing who the source of their belief was and juxtaposing that knowledge with the content of their belief, they came to hold that they knew that there would be an exam sometime next week. Holding that they knew that there would

<sup>(4)</sup> QUINE, *op. cit.*, p. 21, is one of the few to realize that this is so. (Quine actually discusses the execution 'paradox,' but his remarks apply, *mutatis mutandis*, to the case of the surprise examination.)

be an exam sometime next week, knowing that there were five days in a school-week, and having, by hypothesis, eliminated the first four possible exam days, the students arrived at (K). Holding (K), then, they believed that – supposing now that the exam were actually Friday – the exam couldn't be unexpected, since, the antecedent of (K) being affirmed, they'd know there would be an exam beforehand and such knowledge (logically) precludes the exam's being unexpected. Hence, they concluded, no surprise exam Friday. The important point, though, is that they believed (K) because they had been told that there would be an exam next week by an extremely reliable authority.

“But why are both (K) and – supposing once again that the exam were Friday –

(L) I know (or will know) that  
there will be an exam Friday

false?<sup>(5)</sup> If you had said nothing more than ‘There will be an exam sometime next week’ and then waited until Friday, they *would* have known. And the same is true if you had said ‘There will be an exam next Friday.’ But they didn't know because you had said ‘surprise exam.’

“In normal communication circumstances, if a person asserts ‘p’, his listeners believe that p. But if he says ‘p, and you don't believe that p’, or, closer to home, if he says ‘p, and you don't (and won't) expect that p,’ where p is a proposition about the future, then his remark is communicatively anomalous. It's a bit like one of Moore's (so-called) paradoxes, namely, ‘I went to the pictures last night, but I don't believe that I did.’<sup>(6)</sup> (It is *not* like most of the other so-called Pragmatic Paradoxes, e.g., ‘I don't exist’, for there are no circumstances in which such sentences are assertively uttered and the propositions they express true.) Moore's first-person sentence is

<sup>(5)</sup> QUINE, *ibid.*, p. 21, recognizes that (L) is false, but he doesn't note that it is really (K) that is at fault in the students' explicit argument, he doesn't explain why anyone would think either proposition true, he doesn't explain why neither proposition is warranted, and he doesn't explain why both propositions are false.

<sup>(6)</sup> *The Philosophy of G.E. Moore*, ed., P. A. SCHILPP (Evanston, Illinois, 1942), p. 543.

communicatively anomalous, and so is your second-person sentence – though I realize that the indexical ‘you’ could have been replaced in your utterance with some other locution and yet you still could have fooled the students just the way you did. Contextual factors having to do with the circumstances of communication surrounding an utterance, however, do have to be taken into account at some point in order to generate an air of paradox. And in order to explain where the students went wrong, I should add.

“Anyway, ‘You will have an exam next week (next Friday), and you don’t and won’t expect an exam next week (next Friday)’ is communicatively anomalous, because the proposition expressed by the second clause is logically contradictory to a proposition conversationally or communicatively implied (or to use a less recent expression, contextually implied) by the first clause, namely, that those listening to the speaker believe, and so expect, that there will be an exam next week (next Friday). Although the speaker’s compound sentence certainly could express a true conjunction, the fact that it is so grossly communicatively anomalous means that listeners should not rely upon the well-founded principle of communication, or better, the well-founded principle of transmission of knowledge, that usually obtains in communication situations, viz., that listeners are in a position to know, or at least to believe with good reason, the propositions expressed by the speaker’s sentences. Indeed, strict adherence to such principles would mean that the students believe that there would be an exam, believe that the exam would be a total surprise to them, and, since they had been *told* the contents of those beliefs, be conscious of having both of their beliefs. That may be logically possible, but I doubt that it is psychologically possible. In fact, it is this *psychological* impossibility, or at least vast difficulty, which is really or ultimately responsible for the utterance being communicatively anomalous: it is what makes the aforementioned logical contradiction of note and not vice versa. (Moore’s ‘paradox’ should be similarly analyzed, in part, with the speaker and not the listener being the focus of attention.) What the students actually did was first to latch onto ‘There will be an ... exam’, thereby neglecting ‘surprise...’. Doing so, they took the communication situation to be normal, came, through a number of implicit steps, to believe (K), and, assuming that you hadn’t given the exam before Friday, believed (L).



“What I’m saying, in short, is that although

(A) X has been told by a reliable  
authority that p and q

is usually strong evidence for

(B) X knows, or at least has good reason  
for believing, that p and q,

the inference is *not* a good one if q is the proposition that X doesn’t believe (or know, or expect, etc.) that p.

“That’s the main mistake in the students’ argument, and it shows that (K) and (L) are unwarranted. But I said that they were false, not just unwarranted. To show that they are, let me offer the following *reductio ad absurdum* against (L). (K) we needn’t worry about, since we are assuming that the exam *was* given on Friday, and on that assumption, if (L) is false, (K) will be too.

“Assume that

(L) Student Y (‘I’) knows that  
there will be an exam on Friday

is true. (The exam was actually Wednesday, I know, but since I’ve said that the big mistake in the students’ argument has to do with the elimination of Friday, I’m going to continue to concentrate on the Friday exam case.) If true, then

(M) Student Y believes that  
there will be an exam on Friday

would be true also, since knowledge entails belief, or at least entails belief in *this* instance, since student Y is conscious of the fact that he has knowledge. What also follows from (L), of course, is

(N) There will be an exam on  
Friday.

Given (M) and (N), then

(O) Student Y will not be surprised,  
will not be given an exam he did not expect,  
on Friday

would have to be true, too, since belief that an event will occur at certain time precludes surprise at its occurrence at the time of its occurrence. However, student Y *was* surprised when the exam was given out on Friday – remember, the hypothetical Friday is in the same boat as the actual Wednesday – so (O) is false. And since it is false and follows directly from (L), (L), and therefore (K), must be false as well.

“Now, although this *reductio* is sound, it’s probably not satisfying: you would probably still like some sort of explanation of the falsity of (O). Not, I take it, that you question whether (O) is false, for it was as obvious to you as it was to me that the students were surprised. What you probably want to know is *why* the students were surprised when they had consciously said to themselves that there would be an exam and that they knew as much. Strange that certainly is. But not impossible. The explanation is that they had also said to themselves *after* they had come to hold that they knew there would be an exam on Friday, that, according to the teacher, the exam would be unexpected. Reasoning that they knew there would be an exam and that the exam would, supposedly, be unexpected, they deduced that there was a contradiction involved, and so there could be no surprise exam. And, indeed, there *is* a contradiction about, namely

(P) Student Y knows that there  
will be an exam on Friday, and on  
Friday he will be surprised to  
discover that there will be an  
exam that day.

(P) is contradictory because its first conjunct entails that student Y believes that there will be an exam on Friday, and its second conjunct entails that he does not believe that there will be an exam that day. That’s contradictory. Unfortunately, though, all the students confu-

sed (P) with

(Q) There will be a surprise exam  
on Friday,

and probably did that in part because they thought that the first conjunct of (P) was obviously true. Taking (Q) instead of (P) to be contradictory, then, they declared the exam impossible. (Their mistake, in fact, is a bit like Goethe's in declaring his own future non-existence impossible,<sup>(7)</sup> or Berkeley's in declaring material objects impossible:<sup>(8)</sup> they, like Goethe and Berkeley, had put themselves 'in the picture,' illegitimately introduced themselves, as in (P), as part of the *content* of their thoughts.<sup>(9)</sup> Why they took themselves to be knowers, as (P) has them, has already been explained – but I should add that the cases of Goethe and Berkeley are altogether different as far as *that* is concerned.) And here we finally come to the true explanation of why the students were surprised on Friday. They believed that (Q) was contradictory. Believing that, they believed that there could be no exam on Friday. And believing that there could be no exam on Friday, they, naturally enough, were surprised when the exam booklets were handed them. Their own reasoning – and reasoning is a psychological process – had led them to a belief which made possible the exam's being a surprise; hence made possible the falsity of the very belief they had arrived at. That this could, and did, happen is strange, but certainly not paradoxical, in the important sense of the term.

#### IV

"I guess that that about covers where and why the students went wrong, but I suppose you still want to know how I managed to get

(7) See Paul EDWARDS, "'My Death'," *The Encyclopedia of Philosophy*, vol. 5 (New York, 1967), pp. 416-417.

(8) George BERKELEY, *The Principles of Human Knowledge*, section 23.

(9) For an explanation of Goethe's confusion, see EDWARDS, *op. cit.*, pp. 416-417; for a partial explanation of Berkeley's, see Bertrand RUSSELL, *History of Western Philosophy* (London, 1946), pp. 627-628; and for a better explanation, see J.L. MACKIE, *Problems From Locke* (Oxford, 1976), pp. 53-54.

things right. That's a shorter story.

"When you announced the surprise exam, I went through all of the reasoning already detailed and in that way discovered, last week, that a surprise exam was possible, and that if a student rested content with the limited and defective reasoning that all except me did rest content with, then a surprise exam was possible on any day. But I *did not* then say to myself 'He could give us a surprise exam; I have no reason to think he won't. So I guess I should conclude that maybe there will be an exam and maybe there won't be.' Agnosticism of that sort seemed to be the only conclusion warranted at that point;<sup>(10)</sup> but not only would *my* reaching that conclusion make *me*, prudent student that I am, also vulnerable to a surprise exam (a possibility I dislike), but also, and more importantly, I thought I had some reason for thinking that there would be an exam. Let me explain.

"First of all, you are a teacher of philosophy and, as class meetings this semester have shown, not given to making false statements or to speaking carelessly. So I had a good, though not a knockdown, reason for thinking that you wouldn't 'trip' over a statement about a surprise exam. In other words, my past experience had indicated that you wouldn't say such a thing unless you had good reason to think it true. Reinforcing my evidence was the fact that you could make sure that the first conjunct of your statement (parsing it once again as you did, as a conjunction), viz., that there will be an exam, was true easily enough, by simply walking in and giving an exam. And then I realized that although you couldn't make sure that its second conjunct, viz., that the exam would be totally unexpected, was true in any such simple way, you could still make a very rational attempt to make it true, and that that was probably precisely what you were doing. I'll explain that in a moment. For right now all I want to note is that I had some evidence, some very good evidence, for thinking that there would be an exam sometime next week.

"Secondly, I thought that if you didn't give an exam, the students would all say that you had spoken falsely and prove as much just by pointing out that no exam had been given. You then would have been required to give everyone an 'A,' and that, I knew, teachers in general and you in particular don't find particularly agreeable. Since I was sure that you knew as much too, I was pretty sure that there would be an exam.

“Then, taking it that I had good evidence that there would be an exam, but knowing that I didn’t know which day it would be, I gauged my expectations according to the probability that the exam would be administered on a given day. At the beginning of the week, the exam could be on any one of the five days, so the odds were one in five that it would be Monday. So I left a note on your desk saying that those were odds and that my expectations were adjusted accordingly. (Actually, as I explained in my note, they were slightly less than one in five, since I couldn’t be dead certain that there would be an exam; but for simplicity’s sake, I’ll ignore that complication from now on.) No exam Monday. So, with four days left and the odds one in four that the exam would be on any given day, I left another note on your desk before class Tuesday. And similarly on Wednesday, when we actually had the exam. By the end of *last* week, my expectations had all been adjusted to the odds; they were rationally based. I guess you couldn’t say that I *knew* that there would be an exam on Wednesday, since all I really knew was that the odds of that being so were one in three. Still, that didn’t matter, as far as I was concerned. You had glossed ‘surprised’ as ‘having *absolutely no* expectation,’ and I certainly had *some* expectation. That’s how I managed to falsify your statement. And it’s also why I didn’t prepare for the test. I knew that by falsifying your statement and bringing it to your attention, I was guaranteed an ‘A’ on the exam.

“But let me back up a little bit here and explain my earlier remark that you were making a rational attempt to bring about the truth of your second conjunct, viz., that the exam would be unexpected. You uttered certain words, ‘There will be an exam, and the exam will be totally unexpected,’ in a certain context, with listeners who were prudent, knowledgeable, and intelligent – but not geniuses. You knew what your listeners were like. And you also knew that by uttering *those* words in *that* context, you would probably *cause* your listeners to acquire certain beliefs, from which they would probably reason in, well, just the way that, except for me, all did reason. They would then be liable to a surprise exam, and when you actually gave the exam, as you did today, your uttering those words would have been part cause of the very state of affairs that made the (second conjunct of the) proposition the words conveyed true. (In that respect, the situation is like some discussed by William James, e.g., a man who, upon meeting

another man for the first time, says to himself 'That man will be my friend.' His belief, considered now as a psychological state, may well cause him to act in such a way that the stranger does become his friend, that his belief, considered now as a proposition, is true.<sup>(11)</sup> Or, a man on a snowy mountainside might shout 'There will be an avalanche,' and his shouting that might be the cause of the state of affairs that makes the proposition expressed true.)

"There is an irony here. Your compound sentence is a good deal like a Moorean self-defeating sentence, which is conversationally or pragmatically self-undermining. But in the circumstances, your sentence was causally self-verifying, for most students, since its utterance was part cause of the state of affairs that made the proposition expressed true. That irony is, I suppose, about the nearest we get to actual paradox.

"Anyway, you knew all of the above. Since you did want to test us, you said what you did. You rightly thought that most or all students would reason as in fact they did reason. Almost all of them will flunk. But I know you: that wasn't your aim. You were perfectly willing to award an 'A' to any student who could think this way through the thick haze you had created. Any student who reasoned the whole thing out, you thought, well deserved an 'A,' regardless of whether he had prepared for an exam. And you're right about that too: in philosophy it's much more important to be able to reason than to be well prepared for exams. But maybe that's a false dilemma... because, in effect, the *real* surprise exam occurred last week, when you made your announcement."

## V

The instructor was silent for a moment, and then ventured a single comment. "I'm pretty sure you'll be interested in what we'll be doing in class from now on, McQ," he said. "The rest of the semester will be on The Prisoner's Dilemma."

*Philadelphia College of Art*

Michael WREEN

<sup>(10)</sup> QUINE, *op. cit.*, p. 21, strongly suggests that McQ should embrace such agnosticism.

<sup>(11)</sup> William JAMES, "The Will to Believe," as reprinted in William ROWE, ed., *Philosophy of Religion* (New York, 1973), pp. 363-364.