

# Unexpected hanging paradox

From Wikipedia, the free encyclopedia

The **unexpected hanging paradox**, **hangman paradox**, **unexpected exam paradox**, **surprise test paradox** or **prediction paradox** is a paradox about a person's expectations about the timing of a future event (e.g. a prisoner's hanging, or a school test) which he is told will occur at an unexpected time.

Despite significant academic interest, no consensus on its correct resolution has yet been established.<sup>[1]</sup> One approach, offered by the logical school of thought, suggests that the problem arises in a self-contradictory self-referencing statement at the heart of the judge's sentence. Another approach, offered by the epistemological school of thought, suggests the unexpected hanging paradox is an example of an *epistemic paradox* because it turns on our concept of *knowledge*.<sup>[2]</sup> Even though it is apparently simple, the paradox's underlying complexities have even led to it being called a "significant problem" for philosophy.<sup>[3]</sup>

## Contents

- 1 Description of the paradox
- 2 The logical school
  - 2.1 Objections
  - 2.2 Leaky inductive argument
  - 2.3 Additivity of surprise
- 3 The epistemological school
- 4 See also
- 5 References
- 6 External links

## Description of the paradox

The paradox has been described as follows:<sup>[4]</sup>

A judge tells a condemned prisoner that he will be hanged at noon on one weekday in the following week but that the execution will be a surprise to the prisoner. He will not know the day of the hanging until the executioner knocks on his cell door at noon that day.

Having reflected on his sentence, the prisoner draws the conclusion that he will escape from the hanging. His reasoning is in several parts. He begins by concluding that the "surprise hanging" can't be on a Friday, as if he hasn't been hanged by Thursday, there is only one day left - and so it won't be a surprise if he's hanged on a Friday. Since the judge's sentence stipulated that the hanging would be a surprise to him, he concludes it cannot occur on Friday.

He then reasons that the surprise hanging cannot be on Thursday either, because Friday has already been eliminated and if he hasn't been hanged by Wednesday night, the hanging must occur on Thursday, making a Thursday hanging not a surprise either. By similar reasoning he concludes that the hanging can also not occur on Wednesday, Tuesday or Monday. Joyfully he retires to his cell confident that the hanging will not occur at all.

The next week, the executioner knocks on the prisoner's door at noon on Wednesday — which,

despite all the above, will still be an utter surprise to him. Everything the judge said has come true.

Other versions of the paradox replace the death sentence with a surprise fire drill, examination, or lion behind a door or when the bin will be emptied.<sup>[1]</sup>

The informal nature of everyday language allows for multiple interpretations of the paradox. In the extreme case, a prisoner who is paranoid might feel certain in his knowledge that the executioner will arrive at noon on Monday, then certain that he will come on Tuesday and so forth, thus ensuring that every day really is a "surprise" to him. But even without adding this element to the story, the vagueness of the account prohibits one from being objectively clear about which formalization truly captures its essence. There has been considerable debate between the logical school, which uses mathematical language, and the epistemological school, which employs concepts such as knowledge, belief and memory, over which formulation is correct.

## The logical school

Formulation of the judge's announcement into formal logic is made difficult by the vague meaning of the word "surprise". An attempt at formulation might be:

- *The prisoner will be hanged next week and the date (of the hanging) will not be deducible in advance from the assumption that the hanging will occur during the week (A).*

Given this announcement the prisoner can deduce that the hanging will not occur on the last day of the week. However, in order to reproduce the next stage of the argument, which eliminates the penultimate day of the week, the prisoner must argue that his ability to deduce, from statement (A), that the hanging will not occur on the last day, implies that a last-day hanging *would not be surprising*. But since the meaning of "surprising" has been restricted to *not deducible from the assumption that the hanging will occur during the week* instead of *not deducible from statement (A)*, the argument is blocked.

This suggests that a better formulation would in fact be:

- *The prisoner will be hanged next week and its date will not be deducible in advance using this statement as an axiom (B).*

Some authors have claimed that the self-referential nature of this statement is the source of the paradox. Fitch<sup>[5]</sup> has shown that this statement can still be expressed in formal logic. Using an equivalent form of the paradox which reduces the length of the week to just two days, he proved that although self-reference is not illegitimate in all circumstances, it is in this case because the statement is self-contradictory.

## Objections

The first objection often raised to the logical school's approach is that it fails to explain how the judge's announcement appears to be vindicated after the fact. If the judge's statement is self-contradictory, how does he manage to be right all along? This objection rests on an understanding of the conclusion to be that the judge's statement is self-contradictory and therefore the source of the paradox. However, the conclusion is more precisely that *in order for the prisoner to carry out his argument* that the judge's sentence cannot be fulfilled, he must *interpret* the judge's announcement as (B). A reasonable assumption would be that the judge did not intend (B) but that the prisoner misinterprets his words to reach his paradoxical conclusion. The judge's sentence appears to be vindicated afterwards but the statement which is actually shown to be true is that "the prisoner will be *psychologically* surprised by the hanging". This statement in formal logic would not allow the prisoner's argument to be carried out.

A related objection is that the paradox only occurs because the judge tells the prisoner his sentence (rather than

keeping it secret) — which suggests that the act of declaring the sentence is important. Some have argued that since this action is missing from the logical school's approach, it must be an incomplete analysis. But the action is included implicitly. The public utterance of the sentence and its context changes the judge's meaning to something like "there will be a surprise hanging despite my having told you that there will be a surprise hanging". The logical school's approach does implicitly take this into account.

## Leaky inductive argument

The argument that first excludes Friday, and then excludes the last remaining day of the week is an inductive one. The prisoner assumes that *by Thursday* he will know the hanging is due on Friday, but he does not know that before Thursday. By trying to carry an inductive argument backward in time based on a fact known only by Thursday the prisoner may be making an error. The conditional statement "If I reach Thursday afternoon alive then Friday will be the latest possible day for the hanging" does little to reassure the condemned man. The prisoner's argument in any case carries the seeds of its own destruction because if he is right, then he is wrong, and can be hanged any day including Friday.

The counter-argument to this is that in order to claim that a statement will not be a surprise, it is not necessary to predict the truth or falsity of the statement at the time the claim is made, but only to show that such a prediction will become possible in the interim period. It is indeed true that the prisoner does not know on Monday that he will be hanged on Friday, nor that he will still be alive on Thursday. However, he *does* know on Monday, that if the hangman as it turns out knocks on his door on Friday, he will have already have expected that (and been alive to do so) since Thursday night - and thus, if the hanging occurs on Friday then it will certainly have ceased to be a surprise at some point in the interim period between Monday and Friday. The fact that it has not yet ceased to be a surprise at the moment the claim is made is not relevant. This works for the inductive case too. When the prisoner wakes up on any given day, on which the last possible hanging day is *tomorrow*, the prisoner will indeed not know for certain that he will survive to see tomorrow. However, he does know that *if he does* survive today, he will *then* know for certain that he must be hanged tomorrow, and thus by the time he is actually hanged tomorrow it will have ceased to be a surprise. This removes the leak from the argument.

In other words, his reasoning is incorrect, as if the hanging was on Friday, he will have found it unexpected because he would have expected no hanging. It would be true even if the judge said: "You will unexpectedly be hanged today".

## Additivity of surprise

A further objection raised by some commentators is that the property of *being a surprise* may not be additive over cosmophases. For example, the event of "a person's house burning down" would probably be a surprise to him, but the event of "a person's house either burning down or not burning down" would certainly not be a surprise, as one of these must always happen, and thus it is absolutely predictable that the combined event will happen. Which particular one of the combined events actually happens can still be a surprise. By this argument, the prisoner's arguments that each day cannot be a surprise do not follow the regular pattern of induction, because adding extra "non-surprise" days only dilutes the argument rather than strengthening it. By the end, all he has proven is that he will not be surprised to be hanged sometime during the week - but he would not have been anyway, as the judge already told him this in statement (A).

## The epistemological school

Various epistemological formulations have been proposed which show that the prisoner's tacit assumptions about what he will know in the future, together with several plausible assumptions about knowledge, are inconsistent.

Chow (1998) provides a detailed analysis of a version of the paradox in which a surprise examination is to take place on one of two days. Applying Chow's analysis to the case of the unexpected hanging (again with the week shortened to two days for simplicity), we start with the observation that the judge's announcement seems to affirm three things:

- **S1:** *The hanging will occur on Monday or Tuesday.*
- **S2:** *If the hanging occurs on Monday, then the prisoner will not know on Sunday evening that it will occur on Monday.*
- **S3:** *If the hanging occurs on Tuesday, then the prisoner will not know on Monday evening that it will occur on Tuesday.*

As a first step, the prisoner reasons that a scenario in which the hanging occurs on Tuesday is impossible because it leads to a contradiction: on the one hand, by **S3**, the prisoner would not be able to predict the Tuesday hanging on Monday evening; but on the other hand, by **S1** and process of elimination, the prisoner *would* be able to predict the Tuesday hanging on Monday evening.

Chow's analysis points to a subtle flaw in the prisoner's reasoning. What is impossible is not a Tuesday hanging. Rather, what is impossible is a situation in which *the hanging occurs on Tuesday despite the prisoner knowing on Monday evening that the judge's assertions S1, S2, and S3 are all true.*

The prisoner's reasoning, which gives rise to the paradox, is able to get off the ground because the prisoner tacitly assumes that on Monday evening, he will (if he is still alive) know **S1**, **S2**, and **S3** to be true. This assumption seems unwarranted on several different grounds. It may be argued that the judge's pronouncement that something is true can never be sufficient grounds for the prisoner *knowing* that it is true. Further, even if the prisoner knows something to be true in the present moment, unknown psychological factors may erase this knowledge in the future. Finally, Chow suggests that because the statement which the prisoner is supposed to "know" to be true is a statement about his *inability* to "know" certain things, there is reason to believe that the unexpected hanging paradox is simply a more intricate version of Moore's paradox. A suitable analogy can be reached by reducing the length of the week to just one day. Then the judge's sentence becomes: *You will be hanged tomorrow, but you do not know that.*

## See also

- Centipede game, the Nash equilibrium of which uses a similar mechanism as its proof.
- Interesting number paradox
- Crocodile Dilemma
- List of paradoxes

## References

1. <sup>a</sup> <sup>b</sup> T. Y. Chow, "The surprise examination or unexpected hanging paradox," *The American Mathematical Monthly* Jan 1998 [1] (<http://www-math.mit.edu/~tchow/unexpected.pdf>)
2. <sup>a</sup> Stanford Encyclopedia discussion of hanging paradox together with other epistemic paradoxes (<http://plato.stanford.edu/entries/epistemic-paradoxes/>)
3. <sup>a</sup> R. A. Sorensen, *Blindspots*, Clarendon Press, Oxford (1988)
4. <sup>a</sup> "Unexpected Hanging Paradox" (<http://mathworld.wolfram.com/UnexpectedHangingParadox.html>) . Wolfram. <http://mathworld.wolfram.com/UnexpectedHangingParadox.html>.
5. <sup>a</sup> Fitch, F., A Goedelized formulation of the prediction paradox, *Amer. Phil. Quart* 1 (1964), 161–164
  - D. J. O'Connor, "Pragmatic Paradoxes", *Mind* 1948, Vol. 57, pp. 358–9. The first appearance of the paradox in print.

The author claims that certain contingent future tense statements cannot come true.

- M. Scriven, "Paradoxical Announcements", *Mind* 1951, vol. 60, pp. 403–7. The author critiques O'Connor and discovers the paradox as we know it today.
- R. Shaw, "The Unexpected Examination" *Mind* 1958, vol. 67, pp. 382–4. The author claims that the prisoner's premises are self-referring.
- C. Wright and A. Sudbury, "the Paradox of the Unexpected Examination," *Australasian Journal of Philosophy*, 1977, vol. 55, pp. 41–58. The first complete formalization of the paradox, and a proposed solution to it.
- A. Margalit and M. Bar-Hillel, "Expecting the Unexpected", *Philosophia* 1983, vol. 13, pp. 337–44. A history and bibliography of writings on the paradox up to 1983.
- C. S. Chihara, "Olin, Quine, and the Surprise Examination" *Philosophical Studies* 1985, vol. 47, pp. 19–26. The author claims that the prisoner assumes, falsely, that if he knows some proposition, then he also knows that he knows it.
- R. Kirkham, "On Paradoxes and a Surprise Exam," *Philosophia* 1991, vol. 21, pp. 31–51. The author defends and extends Wright and Sudbury's solution. He also updates the history and bibliography of Margalit and Bar-Hillel up to 1991.
- T. Y. Chow, "The surprise examination or unexpected hanging paradox," *The American Mathematical Monthly* Jan 1998 [2] (<http://www-math.mit.edu/~tchow/unexpected.pdf>)
- P. Franceschi, "Une analyse dichotomique du paradoxe de l'examen surprise", *Philosophiques*, 2005, vol. 32-2, 399-421, English translation ([http://philsci-archive.pitt.edu/archive/00004170/01/A\\_Dichotomic\\_Analysis\\_of\\_the\\_Surprise\\_Examination\\_Paradox.pdf](http://philsci-archive.pitt.edu/archive/00004170/01/A_Dichotomic_Analysis_of_the_Surprise_Examination_Paradox.pdf)) .
- M. Gardner, "The Paradox of the Unexpected Hanging", *The Unexpected Hanging and Other \* Mathematical Diversions* 1969. Completely analyzes the paradox and introduces other situations with similar logic.
- W.V.O. Quine, "On a So-called Paradox", *Mind* 1953, vol. 62, pp. 65–66.
- R. A. Sorensen, "Recalcitrant versions of the prediction paradox", *Australasian Journal of Philosophy* 1982, vol. 69, pp. 355–362.

## External links

- Unexpected Hanging: explained using dramatization (<http://thinkingbuddy.uni.cc/pm/index.php/orkut-archive/unexpected-hanging-paradox-solved>)

Retrieved from "[http://en.wikipedia.org/wiki/Unexpected\\_hanging\\_paradox](http://en.wikipedia.org/wiki/Unexpected_hanging_paradox)"

Categories: Paradoxes

---

- This page was last modified on 14 November 2010 at 08:40.
  - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. See Terms of Use for details.
- Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.