

# Free will, determinism, & indeterminism

# First paper assignment

PHIL 20229: Paradoxes

Due: Thursday, March 4 (the Thursday before spring break)

Below is a description of your first paper assignment. As an alternative to doing this assignment, you may come up with your own topic, though you must get the written approval of me or your TA by email first. If you do this, the question that I or your TA approves should be on the first page of your essay. The papers should be about 5 pages in length, double-spaced and with reasonable margins and font.

A late penalty of 3 points per day, including weekends, will be assessed for any papers which are handed in late.

Plagiarism is a serious and growing problem at Notre Dame and other universities. It is your responsibility to acquaint yourselves with the University's honor code, as well as with the philosophy department's guidelines regarding plagiarism. Both are linked from the course web site.

Many of the paradoxes we have discussed have no uncontroversial solution. Examples include: Zeno's paradoxes, McTaggart's proof of the unreality of time, Kant's antinomies, the problem of the statue and the clay, the puzzles of personal identity, the relationship between free will and determinism, and the rule-following paradox. Pick one of these paradoxes, and defend your own view about how the paradox should be resolved. The solution defended can be one of the views discussed in class, or one that you have come up with on your own.

A good paper will clearly explain the paradox and clearly explain why the solution defended is the best available solution. A very good paper will advance arguments which go beyond the arguments discussed explicitly in lectures.

There is no need to include anything more than a very brief introduction, and no need for a "summing up" paragraph at the end. The focus should be on a clear, concise defense of your view on the relevant topic.

There is no need to do any research for this paper. You are encouraged to spend time thinking about the problem for yourself, rather than reading what other people think about it. If you do use outside sources, any sources you read on the topic, whether or not you quote from them, must be cited at the end of the paper. Any citation style is fine so long as it is clear what you are citing.

Our topic today is freedom of the will. More precisely, our topic is the relationship between freedom of the will and determinism, and a cluster of arguments which seem to show that free will is **incompatible with both determinism and indeterminism, and hence impossible.**

To answer understand these arguments, we first need to get clear about what “determinism” means. Here is what van Inwagen says:

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

The example of “rolling back history” as an illustration of what determinism implies.

It is common to use “determinism” as name for the thesis that we have no free will. This is the source of much confusion. **“Determinism” is the name of a thesis about the laws of nature, and that is all. It is not a thesis about free will, or about what we can predict, or anything else.**

## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

Our question is whether determinism is compatible with free will. We now know what “determinism” means; but what does it mean to say that we have free will?

Here no neat definition is possible, since there is disagreement about what, exactly, it takes for an action to be free. But we can offer some helpful paraphrases: to freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

The question of the compatibility of free will and determinism is then: can it ever be the case that choices A and B are open to you, despite the fact that the laws of nature (and the prior state of the universe) are consistent only with you doing A?

The **incompatibilist** says “No.” The **compatibilist** says “Yes.”



## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

## Free will

To freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

The question of the compatibility of free will and determinism is then: can it ever be the case that choices A and B are open to you, despite the fact that the laws of nature (and the prior state of the universe) are consistent only with you doing A?

The **incompatibilist** says “No.” The **compatibilist** says “Yes.”

Many people have a strong initial intuition that free will and determinism are incompatible, and hence that compatibilism must be false.

van Inwagen provides an argument for this conclusion: the **consequence argument**.

## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

## Free will

To freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

van Inwagen provides an argument for this conclusion: the **consequence argument**.

This argument relies on a principle that van Inwagen calls the “no choice principle”:

Suppose that  $p$  and that no one has (or ever had) any choice about whether  $p$ . And suppose also that the following conditional (if-then) statement is true and that no one has (or ever had) any choice about whether it is true: if  $p$ , then  $q$ . It follows from these two suppositions that  $q$  and that no one has (or ever had) any choice about whether  $q$ .

As van Inwagen says, this principle seems intuitively very plausible: “how could I have a choice about something that is an inevitable consequence of something I have no choice about?”

But if this principle is true, we can show -- with the assumption of two other plausible principles -- that free will is inconsistent with determinism.

## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

## Free will

To freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

## The no choice principle

Suppose that  $p$  and that no one has (or ever had) any choice about whether  $p$ . And suppose also that the following conditional (if-then) statement is true and that no one has (or ever had) any choice about whether it is true: if  $p$ , then  $q$ . It follows from these two suppositions that  $q$  and that no one has (or ever had) any choice about whether  $q$ .

Each of the additional principles in van Inwagen's argument says that we have no choice about something.

The first principle is: We have no choice about events which happened in the distant past.

The second principle is: We have no choice about what the laws of nature are.

Putting these principles together, we can construct an argument for the incompatibility of free will and determinism.

## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

## Free will

To freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

To state the consequence argument, let 'DINOSAUR' stand for the state of the universe during some time when dinosaurs roamed the earth, and let 'DECISION' stand for my decision to not sing the Notre Dame fight song 10 minutes from now.

### The no choice principle

If I have no choice about p, and no choice about whether if p, then q, I have no choice about q.

We have no choice about events which happened in the distant past.

We have no choice about what the laws of nature imply.

### Determinism

The laws of nature + the state of the universe at a time determine a unique future. In particular, the laws determine that if DINOSAUR is the case, then so is DECISION.

## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

## Free will

To freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

To state the consequence argument, let 'DINOSAUR' stand for the state of the universe during some time when dinosaurs roamed the earth, and let 'DECISION' stand for my decision to not sing the Notre Dame fight song 10 minutes from now.

### The no choice principle

If I have no choice about p, and no choice about whether if p, then q, I have no choice about q.

We have no choice about events which happened in the distant past.

We have no choice about what the laws of nature imply.

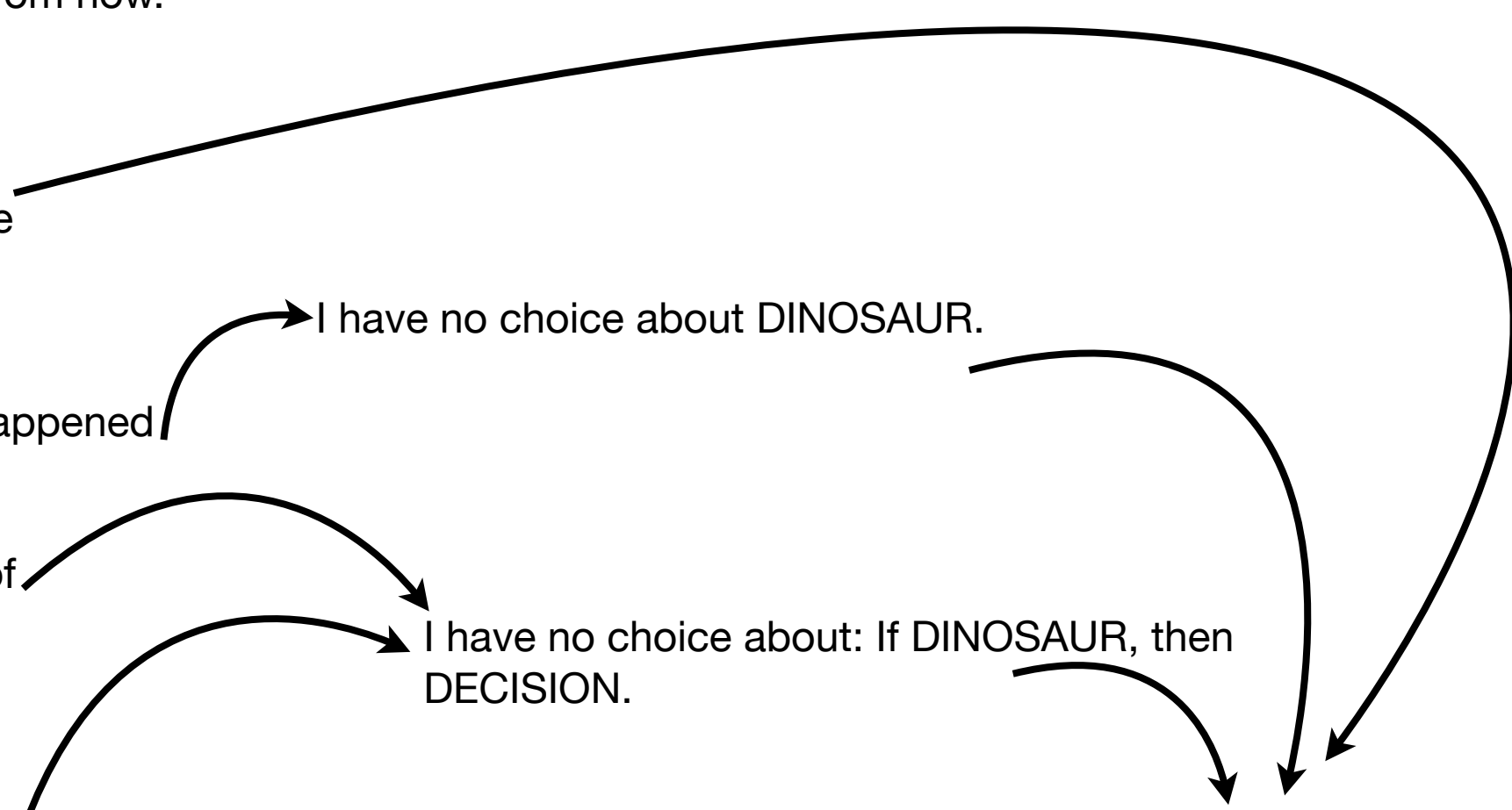
### Determinism

The laws of nature + the state of the universe at a time determine a unique future. In particular, the laws determine that if DINOSAUR is the case, then so is DECISION.

I have no choice about DINOSAUR.

I have no choice about: If DINOSAUR, then DECISION.

I have no choice about DECISION.





## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

### The no choice principle

If I have no choice about p, and no choice about whether if p, then q, I have no choice about q.

We have no choice about events which happened in the distant past.

We have no choice about what the laws of nature imply.

### Determinism

The laws of nature + the state of the universe at a time determine a unique future. In particular, the laws determine that if DINOSAUR is the case, then so is DECISION.

## Free will

To freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

It is important to be clear about the nature of this argument. We begin by assuming the four principles at left, and from them derive that I have no choice about some future action. Since there is nothing special about DECISION, it seems clear that the argument generalizes to show that I have no choice about any of my future actions. So, if we assume these four principles, we can show of any particular action that it is unfree. Hence, if we assume these four principles, we can show that no action is ever free.

If the first three principles are true -- the No Choice Principle + the claims that we have no choice about the laws of nature or the past -- then, what the argument shows is that **if** determinism is true, **then** we have no free will.

## Determinism

Determinism is the thesis that it is true at every moment that the way things then are determines a unique future, that only *one* of the alternative futures that may exist relative to a given moment is a physically possible continuation of the state of things at that moment. Or, if you like, we may say that determinism is the thesis that only one continuation of the state of things at a given moment is consistent with the laws of nature.

### The no choice principle

If I have no choice about p, and no choice about whether if p, then q, I have no choice about q.

We have no choice about events which happened in the distant past.

We have no choice about what the laws of nature imply.

### Determinism

The laws of nature + the state of the universe at a time determine a unique future. In particular, the laws determine that if DINOSAUR is the case, then so is DECISION.

## Free will

To freely choose between A and B is **to be able** to do either of A and B; to freely choose between A and B is for both of A and B **to be open** to you.

If the first three principles are true -- the No Choice Principle + the claims that we have no choice about the laws of nature or the past -- then, what the argument shows is that **if** determinism is true, **then** we have no free will.

So to demonstrate the incompatibility of free will and determinism, we assume the truth of determinism, and argue from there to the absence of free will. This is a style of argument called **conditional proof**. To prove the truth of a statement

if p, then q

we assume p as a premise, and argue from this premise, using only other true premises, to q as our conclusion. If we can construct a valid argument with p + some true statements as premises for q, it follows that the conditional statement

if p, then q

must be true. Here p = the truth of determinism, and q = the denial of the existence of free will.

If the first three principles are true -- the No Choice Principle + the claims that we have no choice about the laws of nature or the past -- then, what the argument shows is that **if** determinism is true, **then** we have no free will.

So to demonstrate the incompatibility of free will and determinism, we assume the truth of determinism, and argue from there to the absence of free will. This is a style of argument called **conditional proof**. To prove the truth of a statement

if p, then q

we assume p as a premise, and argue from this premise, using only other true premises, to q as our conclusion. If we can construct a valid argument with p + some true statements as premises for q, it follows that the conditional statement

if p, then q

must be true. Here p = the truth of determinism, and q = the denial of the existence of free will.

We can lay out the consequence argument in premises as follows:

| The consequence argument |  |   |
|--------------------------|--|---|
| 1                        | The laws of nature + the state of the universe at a time imply a unique future. In particular, (let's suppose) the laws imply that <b>if</b> DINOSAUR is the case, <b>then</b> so is DECISION. | Determinism (premise assumed for conditional proof) |
| 2                        | If the laws of nature imply that X is the case, then we have no choice about X.  | No choice about the laws of nature                  |
| 3                        | We have no choice about the fact that if DINOSAUR, then DECISION.  | 1, 2  |
| 4                        | We have no choice about events which happened in the distant past.   | No choice about the past                            |
| 5                        | We have no choice about DINOSAUR.  | 4   |
| 6                        | If I have no choice about p, and no choice about whether if p, then q, I have no choice about q.   | No choice principle                                 |
| C                        | We have no choice about DECISION.  | 3, 5, 6   |

If premises 2, 4, and 6 are true, then it follows that if Determinism is true, there is no free will. Hence, if these premises are true, then Incompatibilism is true.



### The consequence argument

|   |  |  |
|---|--|--|
| 1 | The laws of nature + the state of the universe at a time imply a unique future. In particular, (let's suppose) the laws imply that <b>if</b> DINOSAUR is the case, <b>then</b> so is DECISION. | Determinism<br>(premise assumed for conditional proof) |
| 2 | If the laws of nature imply that X is the case, then we have no choice about X.  | No choice about the laws of nature                     |
| 3 | We have no choice about the fact that if DINOSAUR, then DECISION.  | 1, 2   |
| 4 | We have no choice about events which happened in the distant past.   | No choice about the past                               |
| 5 | We have no choice about DINOSAUR.  | 4  |
| 6 | If I have no choice about p, and no choice about whether if p, then q, I have no choice about q.   | No choice principle                                    |
| C | We have no choice about DECISION.  | 3, 5, 6  |

So far, this cannot be considered a paradox.

If the consequence argument succeeds, it establishes only the conditional claim that if determinism is true, then we have no free will (and so also that if we have free will, the world must not be deterministic).

One way to think about this is in terms of the following chart illustrating four possible combinations of views about free will and determinism:

|                                    |                               |
|------------------------------------|-------------------------------|
| <del>free will + determinism</del> | no free will + determinism    |
| free will + no determinism         | no free will + no determinism |

In terms of this chart, what the consequence argument establishes, if successful, is that the upper left box does not describe a way the world could be.

It does not, by itself, establish that we do or do not have free will; and it does not by itself establish that determinism is false, or that it is true.

|                               |                                  |
|-------------------------------|----------------------------------|
| free will +<br>determinism    | no free will +<br>determinism    |
| free will + no<br>determinism | no free will + no<br>determinism |

In terms of this chart, what the consequence argument establishes, if successful, is that the upper left box does not describe a way the world could be.

It does not, by itself, establish that we do or do not have free will; and it does not by itself establish that determinism is false, or that it is true.

However, it does limit our choices. One option is that we **deny the existence of free will**; a second option is that we **accept the existence of free will, but say that this is only possible in an indeterministic world**.

The first of these is an extremely difficult option to accept, for at least two reasons.

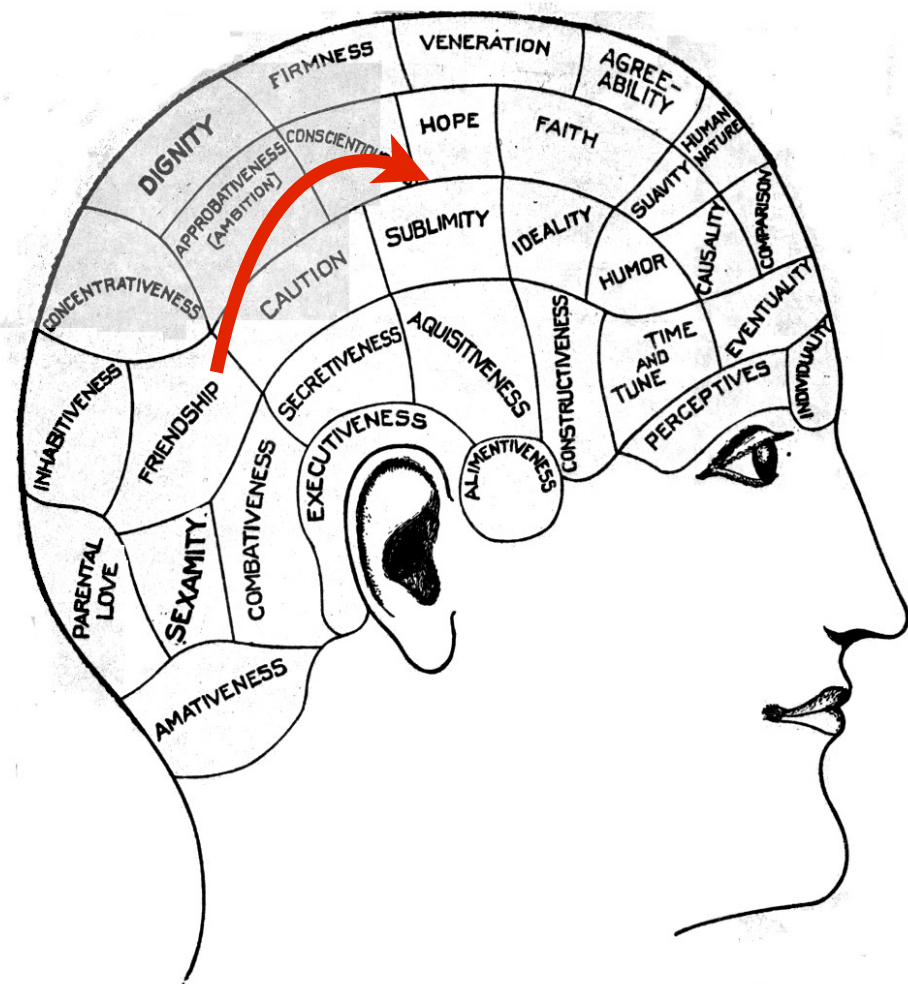
First, it is simply extremely hard to believe that I have no choice about what I will be doing 5 seconds from now. And second, it seems that if there is no free will there is no moral responsibility; and it certainly seems that people are sometimes morally responsible for their actions.

So at this point the view that free will exists, but requires the falsity of determinism, looks like a pretty attractive option.

So at this point the view that free will exists, but requires the falsity of determinism, looks like a pretty attractive option.

However, upon reflection this view too can seem difficult to accept, because it can seem difficult to see how the falsity of determinism could make room for free will.

A good way to introduce the first sort of argument is by way of van Inwagen's example of Jane's decision. We are imagining that Jane is deciding whether to say something, and we are granting for now the incompatibilist's idea that if this decision is free, it must not be determined. Then we can imagine the case going like this:



Make a Mental Picture of Yourself

Each contemplated action will, of course, have antecedents in Jane's cerebral cortex, for it is in that part of Jane (or of her body) that control over her vocal apparatus resides. Let us make a fanciful assumption about these antecedents, since it will make no real difference to our argument what they are. (It will help us to focus our thoughts if we have some sort of mental picture of what goes on inside Jane at the moment of decision.) Let us suppose that there is a certain current-pulse that is proceeding along one of the neural pathways in Jane's brain and that it is about to come to a fork. And let us suppose that if it goes to the left, she will make her confession, and that if it goes to the right, she will remain silent. And let us suppose that it is undetermined which way the pulse will go when it comes to the fork

We now imagine the current pulse traveling through Jane's brain.

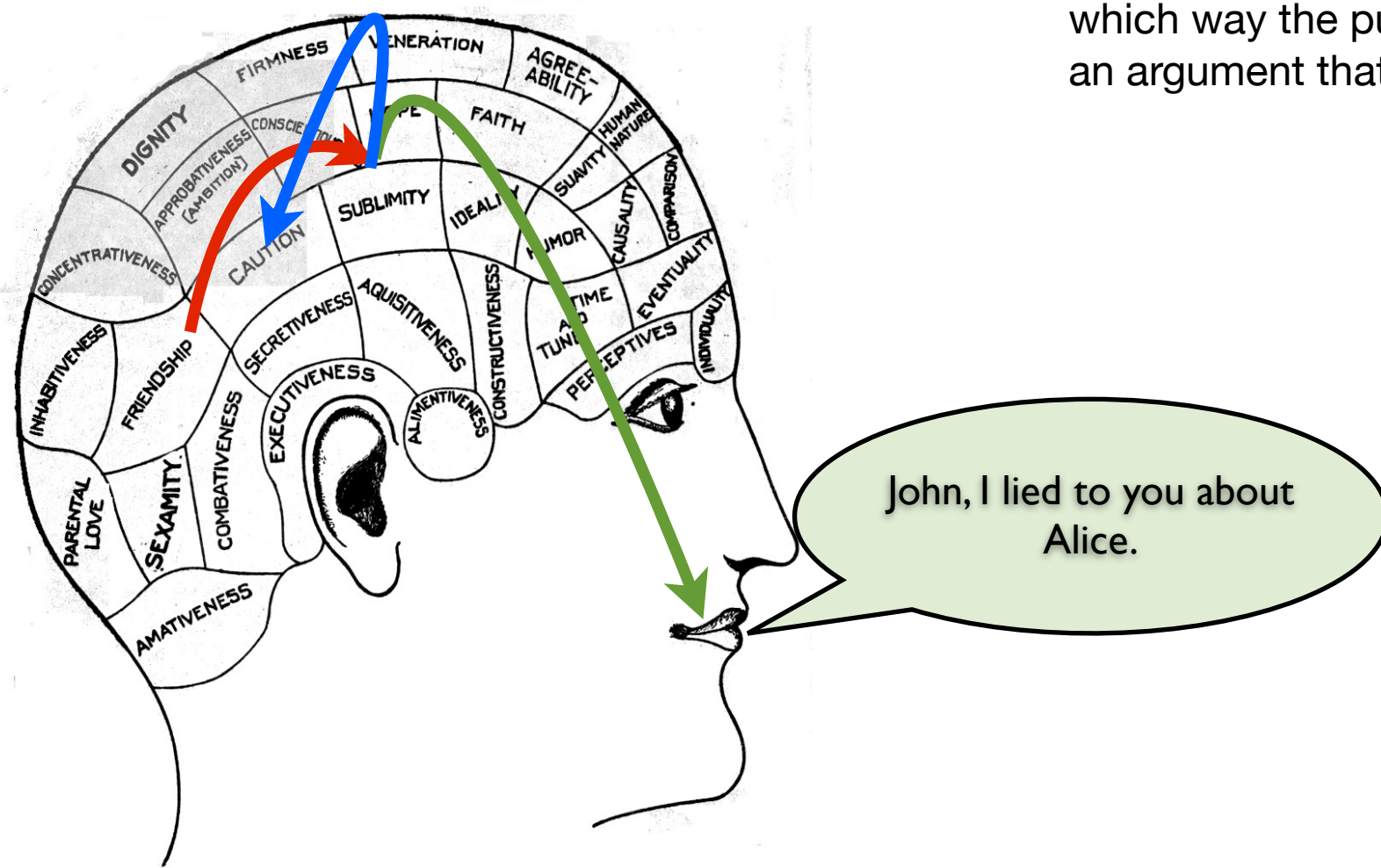
The pulse could go one of two ways. Which way it goes will determine whether or not Jane speaks; and which way will go is not determined by the the laws of nature + the state of Jane's brain (or the state of anything else).

We now imagine the current pulse traveling through Jane's brain.

The pulse could go one of two ways. Which way it goes will determine whether or not Jane speaks; and which way will go is not determined by the laws of nature + the state of Jane's brain (or the state of anything else).

Each contemplated action will, of course, have antecedents in Jane's cerebral cortex, for it is in that part of Jane (or of her body) that control over her vocal apparatus resides. Let us make a fanciful assumption about these antecedents, since it will make no real difference to our argument what they are. (It will help us to focus our thoughts if we have some sort of mental picture of what goes on inside Jane at the moment of decision.) Let us suppose that there is a certain current-pulse that is proceeding along one of the neural pathways in Jane's brain and that it is about to come to a fork. And let us suppose that if it goes to the left, she will make her confession, and that if it goes to the right, she will remain silent. And let us suppose that it is undetermined which way the pulse will go when it comes to the fork

The key question, now is: is Jane free to decide which way the pulse will go? van Inwagen gives an argument that she is not:

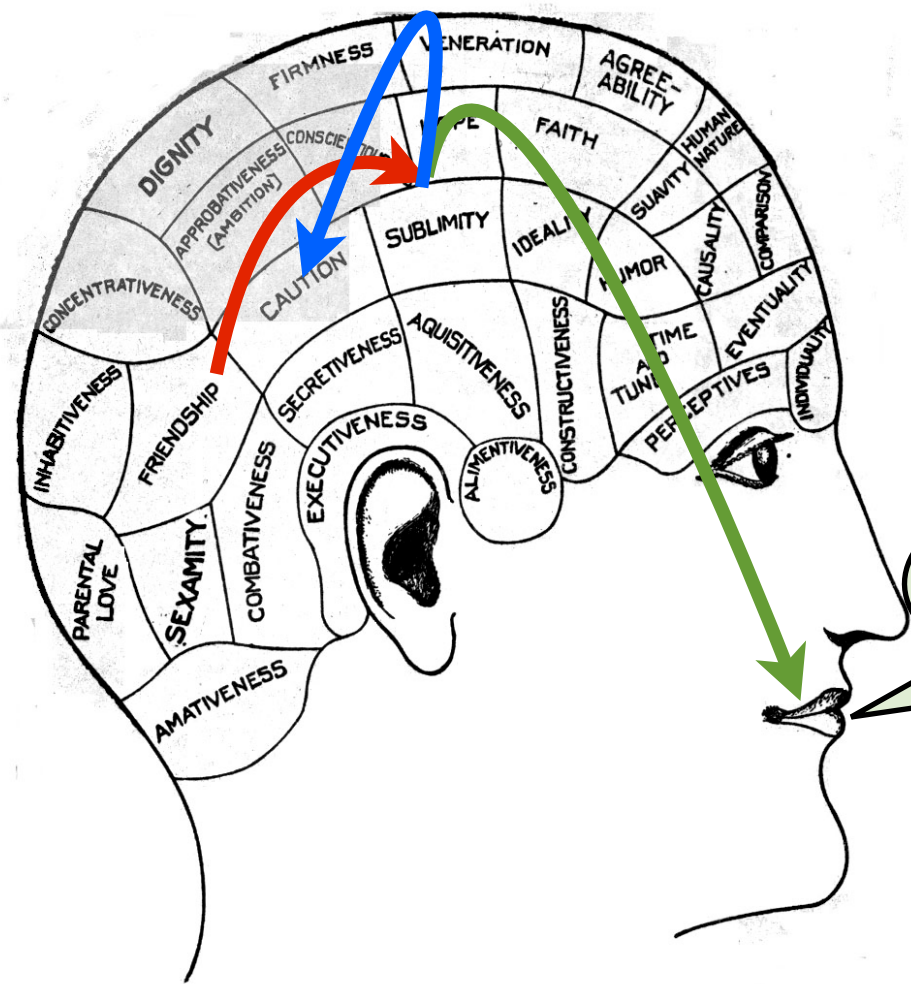


Make a Mental Picture of Yourself



We now imagine the current pulse traveling through Jane's brain.

The pulse could go one of two ways. Which way it goes will determine whether or not Jane speaks; and which way will go is not determined by the laws of nature + the state of Jane's brain (or the state of anything else).



Make a Mental Picture of Yourself

The key question, now is: is Jane free to decide which way the pulse will go? van Inwagen gives an argument that she is not:

Now let us ask: Does Jane have any choice about whether the pulse goes to the left or to the right? If we think about this question for a moment, we shall see that it is very hard to see how she could have any choice about that. Nothing in the way things are at the instant before the pulse makes its "decision" to go one way or the other makes it happen that the pulse goes one way or goes the other. If it goes to the left, that *just happens*. If it goes to the right, *that just happens*. There is no way for Jane to *influence* the pulse. There is no way for her to *make* it go one way rather than the other. Or, at least, there is no way for her to make it go one way rather than the other and leave the "choice" it makes an undetermined event. If Jane did something to make the pulse go to the left, then, obviously, its going to the left would *not* be an undetermined event. It is a plausible idea that the only way to have a choice about the outcome of a process is to be able to arrange things in ways that will make it inevitable that this or that outcome occur. If this plausible idea is right, then it would seem that there is no way in which anyone could have any choice about the outcome of an indeterministic process. And it seems to follow that if, when one is trying to decide what to do, it is truly undetermined what the outcome of one's deliberations will be, then one could have no choice about that outcome.

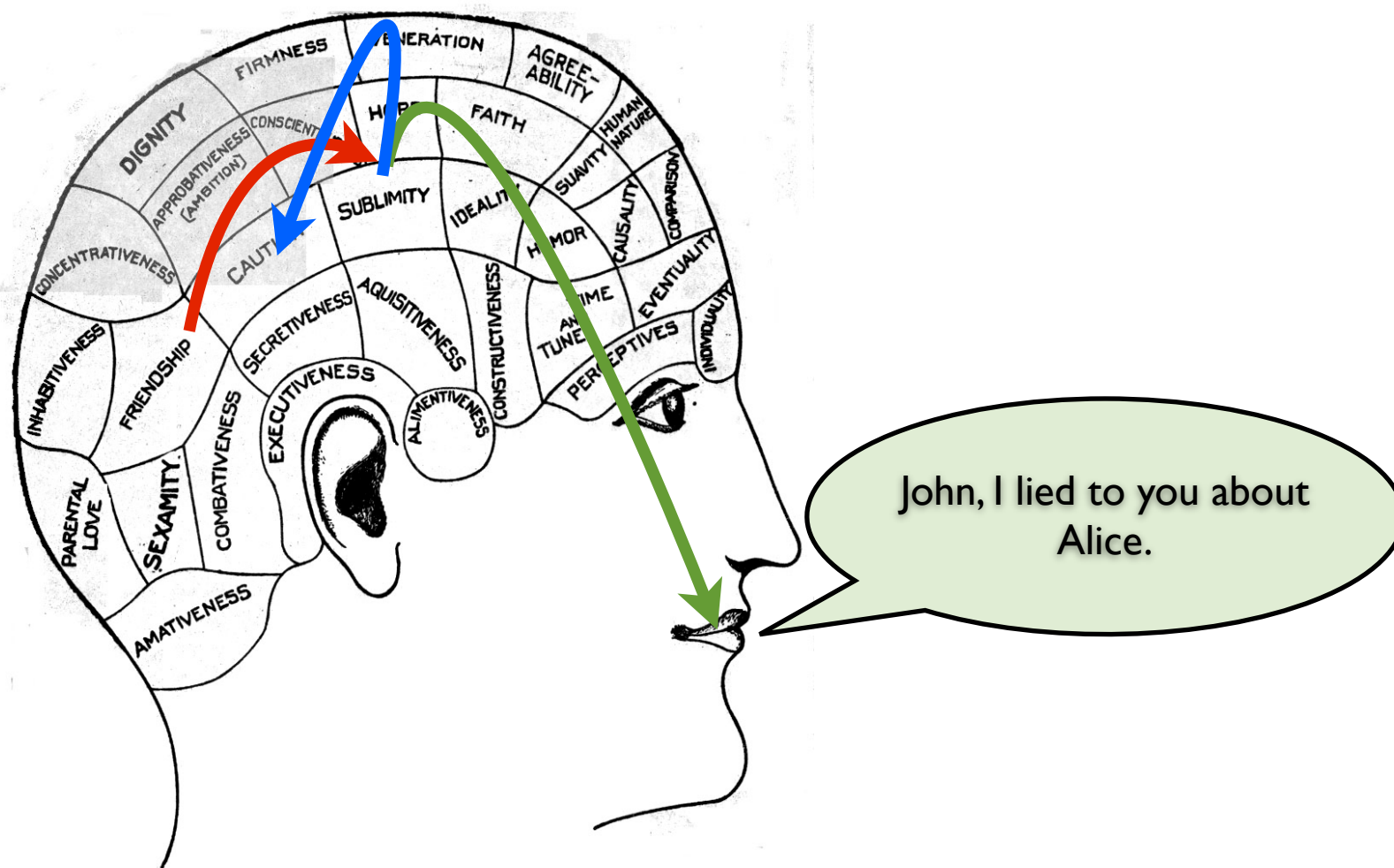
John, I lied to you about Alice.

Now let us ask: Does Jane have any choice about whether the pulse goes to the left or to the right? If we think about this question for a moment, we shall see that it is very hard to see how she could have any choice about that. Nothing in the way things are at the instant before the pulse makes its "decision" to go one way or the other makes it happen that the pulse goes one way or goes the other. If it goes to the left, that *just happens*. If it goes to the right, that just happens. There is no way for Jane to *influence* the pulse. There is no way for her to *make* it go one way rather than the other. Or, at least, there is no way for her to make it go one way rather than the other and leave the "choice" it makes an undetermined event. If Jane did something to make the pulse go to the left, then, obviously, its going to the left would *not* be an undetermined event. It is a plausible idea that the only way to have a choice about the outcome of a process is to be able to arrange things in ways that will make it inevitable that this or that outcome occur. If this plausible idea is right, then it would seem that there is no way in which anyone could have any choice about the outcome of an indeterministic process. And it seems to follow that if, when one is trying to decide what to do, it is truly undetermined what the outcome of one's deliberations will be, then one could have no choice about that outcome.

This argument seems to rely on some principle like the following:

**If nothing determines whether someone chooses A or B, the choice of A or B is random, and hence not a free choice.**

If any principle of this sort is true, this is serious trouble for the incompatibilist who wants to believe in free will. After all, this sort of principle seems to show that free will **requires** determinism -- or at least requires that human actions be determined.



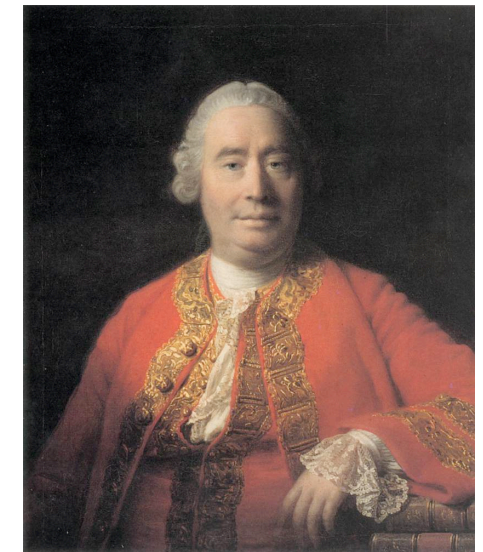


This argument seems to rely on some principle like the following:

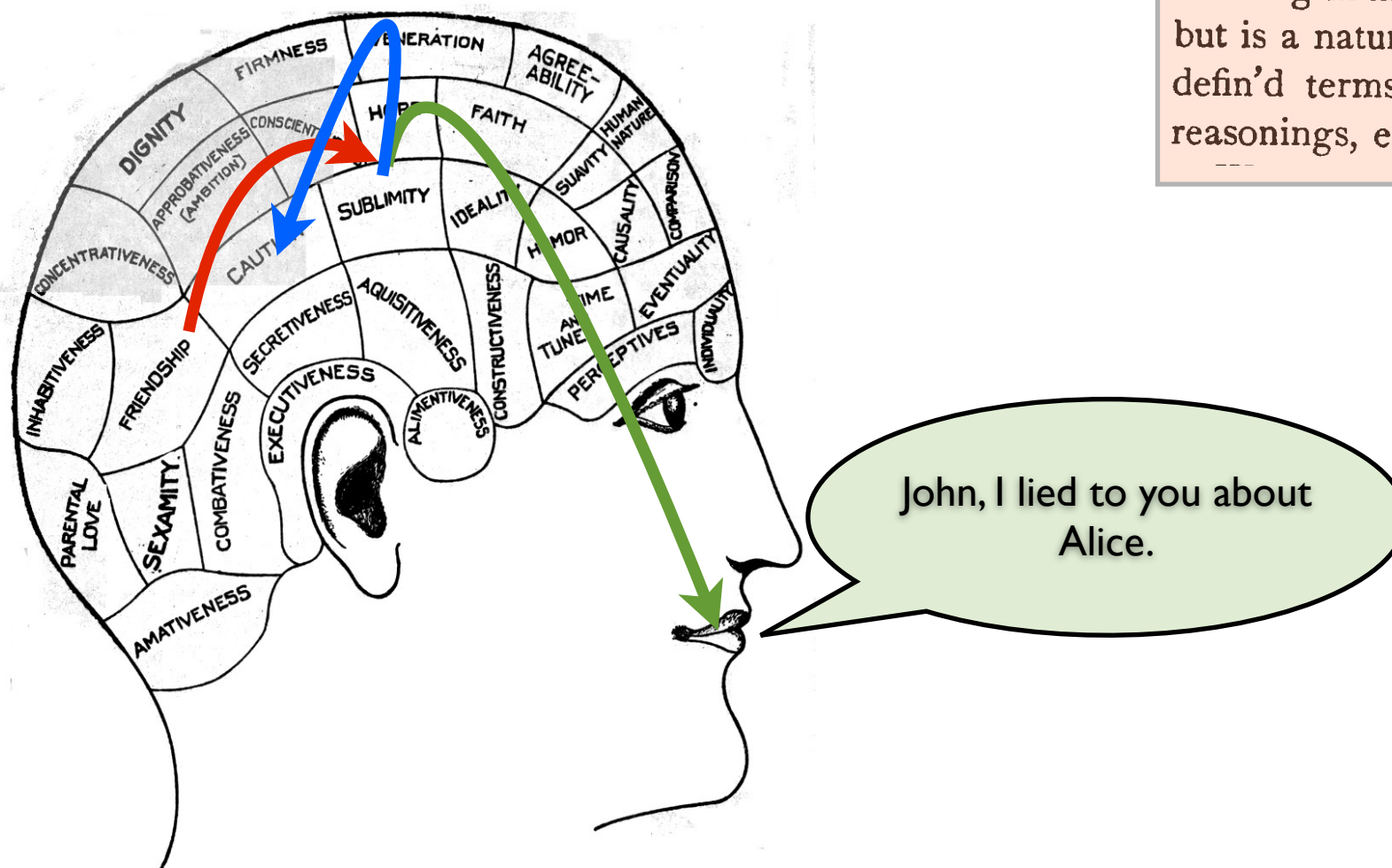
**If nothing determines whether someone chooses A or B, the choice of A or B is random, and hence not a free choice.**

If any principle of this sort is true, this is serious trouble for the incompatibilist who wants to believe in free will. After all, this sort of principle seems to show that free will **requires** determinism -- or at least requires that human actions be determined.

Something like this was the view of David Hume. He thought that genuine free will requires that one's actions be determined by one's prior motives, and that the idea that there is a conflict between determinism and free will just rests on a confusion:



'Tis commonly allow'd that mad-men have no liberty. But were we to judge by their actions, these have less regularity and constancy than the actions of wise-men, and consequently are farther remov'd from necessity. Our way of thinking in this particular is, therefore, absolutely inconsistent; but is a natural consequence of these confus'd ideas and undefin'd terms, which we so commonly make use of in our reasonings, especially on the present subject.



Make a Mental Picture of Yourself

This was part of Hume's argument for **compatibilism**: the view that determinism is consistent with the existence of free will. But, as van Inwagen says, one might take the example of Jane and her brain to show something quite different.

This was part of Hume's argument for **compatibilism**: the view that determinism is consistent with the existence of free will. But, as van Inwagen says, one might take the example of Jane and her brain to show something quite different.

But now a disquieting possibility suggests itself. Perhaps the explanation of the fact that both compatibilism and incompatibilism seem to lead to mysteries is simply that the concept of free will is self-contradictory. Perhaps free will is, as the incompatibilists say, incompatible with determinism. But perhaps it is also incompatible with indeterminism, owing to the impossibility of anyone's having a choice about the outcome of an indeterministic process. If free will is incompatible with both determinism and indeterminism, then, since either determinism or indeterminism has to be true, free will is impossible.

Before we accept this conclusion, we should try to get a bit more explicit about how the argument against incompatibilist free will is supposed to work.

The basic idea seems to be this: if we think about the causal chain leading up to some putatively free action A of Jane's, then, if A is really free and incompatibilism is true, there must be some event, E, in this causal chain which is not determined by prior events plus the laws of nature. Further, it seems that for A to be free, Jane must have had a choice about whether E happened. But it is hard to see how Jane could have had a choice about whether E happened, since the entire state of the universe prior to E, including everything Jane does and thinks, is consistent both with E happening and with E not happening. But then it was **not up to Jane** whether E happened at all.

One way to bring out the intuition here - that if the history of the universe up to E is consistent with E happening and E not happening, it can't be up to Jane whether E happens - is brought out nicely by a thought experiment which van Inwagen discusses in his paper "Free will remains a mystery".



One way to bring out the intuition here - that if the history of the universe up to E is consistent with E happening and E not happening, it can't be up to Jane whether E happens - is brought out nicely by a thought experiment which van Inwagen discusses in his paper "Free will remains a mystery".

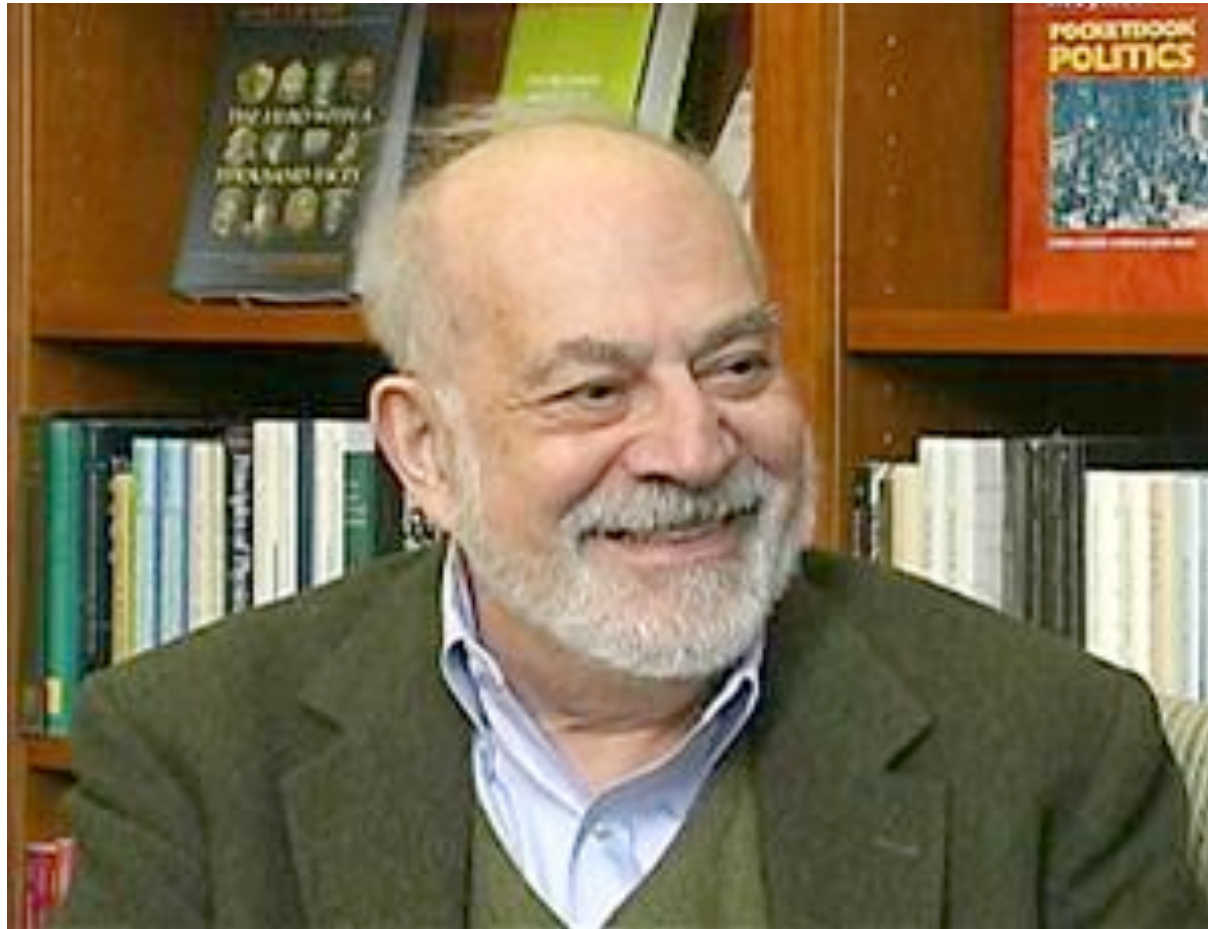
Let us suppose undetermined free acts occur. Suppose, for example, that in some difficult situation Alice was faced with a choice between lying and telling the truth and that she freely chose to tell the truth—or, what is the same thing, she seriously considered telling the truth, seriously considering lying, told the truth, and was able to tell the lie she had been contemplating. And let us assume that free will is incompatible with determinism, and that Alice's telling the truth, being a free act, was therefore undetermined. Now suppose that immediately after Alice told the truth, God caused the universe to revert to precisely its state one minute before Alice told the truth (let us call the first moment the universe was in this state ' $t_1$ ' and the second moment the universe was in this state ' $t_2$ '), and then let things "go forward again." What would have happened the second time? What would have happened after  $t_2$ ? Would she have lied or would she have told the truth? Since Alice's "original" decision, her decision to tell the truth, was undetermined—since it was undetermined whether she would lie or tell the truth—, her "second" decision would also be undetermined, and this question can therefore have no answer; or it can have no answer but, "Well, although she would either have told the truth or lied, it's not the case that she would have told the truth and it's not the case that she would have lied; lying is not what she would have done, and telling the truth is not what she would have done. One can say only that she *might* have lied and she *might* have told the truth."

Now let us suppose that God *a thousand times* caused the universe to revert to exactly the state it was in at  $t_1$  (and let us suppose that we are somehow suitably placed, metaphysically speaking, to observe the whole sequence of "replays"). What would have happened? What should we expect to observe? Well, again, we can't say what would have happened, but we can say what would *probably* have happened: sometimes Alice would have lied and sometimes she would have told the truth. As the number of "replays" increases, we observers shall—almost certainly—observe the ratio of the outcome "truth" to the outcome "lie" settling down to, converging on, some value.<sup>16</sup> We may, for example, observe that, after a fairly large number of replays, Alice lies in thirty percent of the replays and tells the truth in seventy percent of them—and that the figures 'thirty percent' and 'seventy percent' become more and more accurate as the number of replays increases. But let us imagine the simplest case: we observe that Alice tells the truth in about half the replays and lies in about half the replays. If, after one hundred replays, Alice has told the truth fifty-three times and has lied forty-eight times,<sup>17</sup> we'd begin strongly to suspect that the

figures after a thousand replays would look something like this: Alice has told the truth four hundred and ninety-three times and has lied five hundred and eight times. Let us suppose that these are indeed the figures after a thousand replays. Is it not true that as we watch the number of replays increase, we shall become convinced that what will happen in the *next* replay is a matter of chance? (The compulsive gamblers among us might find themselves offering bets about what Alice would do in the next replay.) If we have watched seven hundred and twenty-six replays, we shall be faced with the inescapable impression that what happens in the seven-hundred-and-twenty-seventh replay will be due simply to chance. Is there any reason we should resist this impression? Well, we certainly know that there is nothing we could learn about the situation that could undermine the impression, for we already know everything that is relevant to evaluating it: we know that the outcome of the seven-hundred-and-twenty-seventh replay will not be determined by its initial state (the common initial state of all the replays) and the laws of nature. Each time God places the universe in this state, both "truth" and "lie" are consistent with the universe's being in this state and the laws of nature. A sheaf of possible futures (possible in the sense of being consistent with the laws) leads "away" from this state, and, if the sheaf is assigned a measure of 1, surely, we must assign a measure of 0.5 to the largest sub-sheaf in all of whose members Alice tells the truth and the same measure to the largest sub-sheaf in all of whose members she lies. We must make this assignment because it is the only reasonable explanation of the observed approximate equality of the "truth" and "lie" outcomes in the series of replays. And if we accept this general conclusion, what other conclusion can we accept about the seven-hundred-and-twenty-seventh replay (which is about to commence) than this: each of the two possible outcomes of this replay has an objective, "ground-floor" probability of 0.5—and there's nothing more to be said? And this, surely, means that, in the strictest sense imaginable, the outcome of the replay will be a matter of chance.

Now, obviously, what holds for the seven-hundred-and-twenty-seventh replay holds for all of them, including the one that wasn't strictly a *replay*, the initial sequence of events. But this result concerning the "initial replay", the "play," so to speak, should hold whether or not God bothers to produce any replays. And if He does not—well, that's just the actual situation. Therefore, an undetermined action is simply a matter of chance: if it was undetermined in the one, actual case whether Alice lied or told the truth, it was a mere matter of chance whether she lied or told the truth.

So one problem for incompatibilism + belief in free will is the apparent link between undetermined actions and random actions. A second problem is brought out by an important example due to Harry Frankfurt.



“Suppose someone --- Black, let us say --- wants Jones to perform a certain action. Black is prepared to go to considerable lengths to get his way, but he prefers to avoid showing his hand unnecessarily. So he waits until Jones is about to make up his mind what to do, and does nothing unless it is clear to him (Black is an excellent judge of such things) that Jones is going to decide to do something *other* than what he wants him to do. If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do....

Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it. It would be quite unreasonable to excuse Jones for his action ... on the basis of the fact that he could not have done otherwise. This fact played no role at all in leading him to act as he did.... Indeed, everything happened just as it would have happened without Black's presence in the situation and without his readiness to intrude into it.”

“Suppose someone --- Black, let us say --- wants Jones to perform a certain action. Black is prepared to go to considerable lengths to get his way, but he prefers to avoid showing his hand unnecessarily. So he waits until Jones is about to make up his mind what to do, and does nothing unless it is clear to him (Black is an excellent judge of such things) that Jones is going to decide to do something *other* than what he wants him to do. If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do....

Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it. It would be quite unreasonable to excuse Jones for his action ... on the basis of the fact that he could not have done otherwise. This fact played no role at all in leading him to act as he did.... Indeed, everything happened just as it would have happened without Black's presence in the situation and without his readiness to intrude into it.”

Suppose, at time  $T$ , that Black decides that he wants Jones on a bus out of South Bend by some later time,  $T+3$ . He hopes that Jones will get on the bus of his own accord, but, if he doesn't, plans to force him onto the bus.

Now suppose that, at time  $T+1$ , Jones is deliberating about whether or not to take a bus out of South Bend. He goes back and forth, but eventually decides to board the bus.

At time  $T+2$ , Jones boards a bus leaving South Bend. This seems (certainly, at least, to Jones himself) to be a free action. Since it never got to time  $T+3$ , Black never had to execute his nefarious plan.

But now think about Jones' decision making at time  $T+1$ . At that time, it was already determined that Jones would be boarding a bus out of South Bend. After all, Jones must either decide to board the bus, or not. In the former case he gets on the bus, and in the latter case, Black forces him on the bus, so again he gets on the bus. So facts fixed prior to his decision making process determined that he would get on the bus. Nonetheless, his decision to get on the bus seems to be a free action.

This seems to be some sort of evidence that free actions can coexist with determinism.



“Suppose someone --- Black, let us say --- wants Jones to perform a certain action. Black is prepared to go to considerable lengths to get his way, but he prefers to avoid showing his hand unnecessarily. So he waits until Jones is about to make up his mind what to do, and does nothing unless it is clear to him (Black is an excellent judge of such things) that Jones is going to decide to do something *other* than what he wants him to do. If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do....

Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it. It would be quite unreasonable to excuse Jones for his action ... on the basis of the fact that he could not have done otherwise. This fact played no role at all in leading him to act as he did.... Indeed, everything happened just as it would have happened without Black's presence in the situation and without his readiness to intrude into it.”

This seems to be some sort of evidence that free actions can coexist with determinism.

But one might also view the argument another way - especially if one finds the consequence argument for incompatibilism convincing. One might think that what this argument shows is that **if** any of our actual actions are free, then free will is possible even if determinism is true - after all, Jones seems to be free in whatever sense we are free. But then if we think that it is **not** possible for free will and determinism to coexist, it will follow that none of our actual actions are free.

This is why the problem of free will seems so hard: no matter what view one takes of the relationship between free will and determinism, one faces some difficult challenges.

And each of these options, as mentioned, has its costs.

Basically, we have three options:

**Deny that there is any such thing as free will.**

**Affirm the existence of free will, and say that this is compatible with determinism being true.**

**Affirm the existence of free will, and say that this is incompatible with determinism being true.**

This is why the problem of free will seems so hard: no matter what view one takes of the relationship between free will and determinism, one faces some difficult challenges.

And each of these options, as mentioned, has its costs.

Basically, we have three options:

**Deny that there is any such thing as free will.**

Make sense of the fact that it seems so clear that what we will do in the near future is something that we have a choice about.

**Affirm the existence of free will, and say that this is compatible with determinism being true.**

Either deny that anyone is ever morally responsible for anything, or explain how moral responsibility is possible without free will

**Affirm the existence of free will, and say that this is incompatible with determinism being true.**

This is why the problem of free will seems so hard: no matter what view one takes of the relationship between free will and determinism, one faces some difficult challenges.

And each of these options, as mentioned, has its costs.

Basically, we have three options:

**Deny that there is any such thing as free will.**

Make sense of the fact that it seems so clear that what we will do in the near future is something that we have a choice about.

**Affirm the existence of free will, and say that this is compatible with determinism being true.**

Either deny that anyone is ever morally responsible for anything, or explain how moral responsibility is possible without free will

**Affirm the existence of free will, and say that this is incompatible with determinism being true.**

Explain how it can be that I have no choice about p, and no choice about the fact that if p, then q, and yet have a choice about q.

This is why the problem of free will seems so hard: no matter what view one takes of the relationship between free will and determinism, one faces some difficult challenges.

And each of these options, as mentioned, has its costs.

Basically, we have three options:

**Deny that there is any such thing as free will.**

Make sense of the fact that it seems so clear that what we will do in the near future is something that we have a choice about.

**Affirm the existence of free will, and say that this is compatible with determinism being true.**

Either deny that anyone is ever morally responsible for anything, or explain how moral responsibility is possible without free will

**Affirm the existence of free will, and say that this is incompatible with determinism being true.**

Explain how it can be that I have no choice about p, and no choice about the fact that if p, then q, and yet have a choice about q.

Explain how an action can be undetermined without being random, and hence not free.

Make sense of Frankfurt's example; either explain why Jones is really not free, or why Jones's being free is not a genuine example of a free action determined by factors outside of the agent's control.



**Deny that there is any such thing as free will.**

Make sense of the fact that it seems so clear that what we will do in the near future is something that we have a choice about.

Either deny that anyone is ever morally responsible for anything, or explain how moral responsibility is possible without free will

**Affirm the existence of free will, and say that this is compatible with determinism being true.**

Explain how it can be that I have no choice about p, and no choice about the fact that if p, then q, and yet have a choice about q.

**Affirm the existence of free will, and say that this is incompatible with determinism being true.**

Explain how an action can be undetermined without being random, and hence not free.

Make sense of Frankfurt's example; either explain why Jones is really not free, or why Jones's being free is not a genuine example of a free action determined by factors outside of the agent's control.

This is why free will seems to be such a great mystery. It is clear that one of the views listed at left must be true, but it is hard to see how any of them could be.