

# Philosophical Argumentation

Jan Willem Wieland  
[j.j.w.wieland@vu.nl](mailto:j.j.w.wieland@vu.nl)

2015 - updated 1/2/21

## Synopsis

Consider typical philosophical claims: that moral responsibility does not require that we could have acted otherwise, that knowledge should consist of more than a belief that is true and justified, that reality is to be preferred above happiness, and so on. The question is: how do philosophers actually defend such statements? And how can we judge whether they are doing it well? This textbook addresses the basic argumentative skills of the philosopher.

## Skills:

1. Identifying standpoints
2. Analysing arguments
3. Reading texts
4. Evaluating arguments as to their validity
5. Analysing implications
6. Translating into propositional logic
7. Making deductions in propositional logic
8. Evaluating arguments as to their soundness
9. Unmasking fallacies
10. Rewriting existing texts
11. Recognising counterexamples
12. Formulating definitions
13. Analysing thought experiments
14. Reasoning with thought experiments
15. Determining reflective equilibrium
16. Formulating standpoints
17. Writing arguments
18. Presenting and giving feedback
19. Finding literature
20. Taking the exam

## 1. Standpoints

Philosophical texts are a special kind of text. They are texts that attempt to prove something, texts that make a certain argument.

Our task will be to evaluate these arguments (and to write down our own arguments later).

We will do this in different steps. The first step is to identify the standpoint of the author, what the author wants to prove.

This is a crucial skill: if you have discovered that standpoint, you can more easily follow the author's thinking and understand the text.

Ideally, the text begins or ends with articulating this standpoint:

“Property is sacred, no one may take from me what I possess. This is something that many of us tacitly accept, but I will show that it is mistaken. ...”

“... Hence, I have shown that it is incorrect to claim that no one may take from me what I possess. Property is not sacred.”

The argument for the standpoint (here: “property is not sacred”) is developed in the rest of the text. Still, standpoints are sometimes more difficult to find than readers would like.

//

Assignment 1.1:

Identify the main standpoint in the following text by Williams:<sup>1</sup>

[L]et us look ... at two examples, to see what utilitarianism might say about them, what we might say about utilitarianism and, most importantly of all, what would be implied by certain ways of thinking about the situations...

(1) George, who has just taken his Ph.D. in chemistry, finds it extremely difficult to get a job. He is not very robust in health, which cuts down the number of jobs he might be able to do satisfactorily. His wife has to go out to work to keep them, which itself causes a great deal of strain, since they have small children and there are severe problems about looking after them. The results of this, especially on the children, are damaging. An older chemist, who knows about this situation, says that he can get George a decently paid job in a certain laboratory, which pursues research into chemical and biological warfare. George says that he cannot accept this, since he is opposed to chemical and biological warfare. The older man replies that he is not too keen on it himself, come to that, but after all George's refusal is not going to make the job or the laboratory go away; what is more, he happens to know that if George refuses the job, it will certainly go to a contemporary of

---

<sup>1</sup> Williams, B. 1973. A Critique of Utilitarianism. In *Utilitarianism. For and Against*. CUP. Note: you can find all references in this textbook on your own: see §19.

George's who is not inhibited by any such scruples and is likely if appointed to push along the research with greater zeal than George would. Indeed, it is not merely concern for George and his family, but (to speak frankly and in confidence) some alarm about this other man's excess of zeal, which has led the older man to offer to use his influence to get George the job... George's wife, to whom he is deeply attached, has views (the details of which need not concern us) from which it follows that at least there is nothing particularly wrong with research into CBW. What should he do?

(2) Jim finds himself in the central square of a small South American town. Tied up against the wall are a row of twenty Indians, most terrified, a few defiant, in front of them several armed men in uniform. A heavy man in a sweat-stained khaki shirt turns out to be the captain in charge and, after a good deal of questioning of Jim which establishes that he got there by accident while on a botanical expedition, explains that the Indians are a random group of the inhabitants who, after recent acts of protest against the government, are just about to be killed to remind other possible protestors of the advantages of not protesting. However, since Jim is an honoured visitor from another land, the captain is happy to offer him a guest's privilege of killing one of the Indians himself. If Jim accepts, then as a special mark of the occasion, the other Indians will be let off. Of course, if Jim refuses, then there is no special occasion, and Pedro here will do what he was about to do when Jim arrived, and kill them all. Jim, with some desperate recollection of schoolboy fiction, wonders whether if he got hold of a gun, he could hold the captain, Pedro and the rest of the soldiers to threat, but it is quite clear from the set-up that nothing of that kind is going to work: any attempt at that sort of thing will mean that all the Indians will be killed, and himself. The men against the wall, and the other villagers, understand the situation, and are obviously begging him to accept. What should he do?

To these dilemmas, it seems to me that utilitarianism replies, in the first case, that George should accept the job, and in the second, that Jim should kill the Indian. Not only does utilitarianism give these answers but, if the situations are essentially as described and there are no further special factors, it regards them, it seems to me, as obviously the right answers. But many of us would certainly wonder whether, in (1), that could possibly be the right answer at all; and in the case of (2), even one who came to think that perhaps that was the answer, might well wonder whether it was obviously the answer. Nor is it just a question of the rightness or obviousness of these answers. It is also a question of what sort of considerations come into finding the answer. A feature of utilitarianism is that it cuts out a kind of consideration which for some others makes a difference to what they feel about such cases: a consideration involving the idea, as we might first and very simply put it, that each of us is specially responsible for what he does, rather than for what other people do. This is an idea closely connected with the value of integrity. It is often suspected that utilitarianism, at least in its direct forms, makes integrity as a value more or less unintelligible. I shall try to show that this suspicion is correct. Of course, even if that is correct, it would not necessarily follow that we should reject utilitarianism; perhaps, as utilitarians sometimes suggest, we should just forget about integrity, in favour of such things as a concern for

the general good. However, if I am right, we cannot merely do that, since the reason why utilitarianism cannot understand integrity is that it cannot coherently describe the relations between a man's projects and his actions. (1973: 96-100)

//

Assignment 1.2:

Identify the main standpoint in the following text by Singer:<sup>2</sup>

I begin with the assumption that suffering and death from lack of food, shelter, and medical care are bad. I think most people will agree about this, although one may reach the same view by different routes. I shall not argue for this view. People can hold all sorts of eccentric positions, and perhaps from some of them it would not follow that death by starvation is in itself bad. It is difficult, perhaps impossible, to refute such positions, and so for brevity I will henceforth take this assumption as accepted. Those who disagree need read no further. My next point is this: if it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it. By "without sacrificing anything of comparable moral importance" I mean without causing anything else comparably bad to happen, or doing something that is wrong in itself, or failing to promote some moral good, comparable in significance to the bad thing that we can prevent. This principle seems almost as uncontroversial as the last one. It requires us only to prevent what is bad, and not to promote what is good, and it requires this of us only when we can do it without sacrificing anything that is, from the moral point of view, comparably important. I could even, as far as the application of my argument to the Bengal emergency is concerned, qualify the point so as to make it: if it is in our power to prevent something very bad from happening, without thereby sacrificing anything morally significant, we ought, morally, to do it. An application of this principle would be as follows: if I am walking past a shallow pond and see a child drowning in it, I ought to wade in and pull the child out. This will mean getting my clothes muddy, but this is insignificant, while the death of the child would presumably be a very bad thing.

The uncontroversial appearance of the principle just stated is deceptive. If it were acted upon, even in its qualified form, our lives, our society, and our world would be fundamentally changed. For the principle takes, firstly, no account of proximity or distance. It makes no moral difference whether the person I can help is a neighbor's child ten yards from me or a Bengali whose name I shall never know, ten thousand miles away. Secondly, the principle makes no distinction between cases in which I am the only person who could possibly do anything and cases in which I am just one among millions in the same position. (1972: 231-2)

Singer's standpoint:

---

<sup>2</sup> Singer, P. 1972. Famine, Affluence, and Morality. *Philosophy & Public Affairs* 1.

“If it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it.”

There is no clear argument for this standpoint in the above fragment (other than “This principle seems almost as uncontroversial as the last one”).

Nevertheless, the example of the pond can be used and generalised:

- (1) If you can save the child by jumping into the pond without thereby having to sacrifice something of comparable moral importance, then you should jump into the pond.
- (2) What holds in the case of the pond holds in general.
- (3) Therefore: if S can prevent something bad by doing A, without having thereby to sacrifice something of comparable moral importance, then S should do A.<sup>3</sup>

You can already see that such reconstructions require some creativity! After all, this argument cannot be found in this way in Singer’s text.

Such arguments (1)-(3) consist of premises and a conclusion.

Premise: a proposition that is considered to be true/plausible in an argument and is used to support a conclusion.

The premises here are (1) and (2).

Conclusion (or inference): a proposition that is derived from one or more propositions.

The conclusion here is (3).<sup>4</sup>

Conclusions often follow “therefore” and premises precede it:  
[Premises] therefore [conclusion].

But the order can also be reversed:  
[Conclusion] because [premises].

Note: not all instances of “because” are instances of arguments. Consider: “I do philosophy because I like it.” That I like it constitutes a reason for me to do philosophy, no reason *to believe* that I do philosophy.

Important:

You may disagree with Singer’s premises (1)-(2) and/or conclusion (3). For the time being, though, we will focus on *reconstruction* and look at *evaluation* later.

Assignment 1.3:

Read Singer’s text in its entirety and identify further (main) standpoints he also defends. Also, indicate how they are connected.

---

<sup>3</sup> “S” stands for any individual agent and “A” an act.

<sup>4</sup> Later in §4 we will ask if inferences like (3) may actually be made. How does (3) follow from (1) and (2)?

## 2. Arguments

In reconstructions of arguments all premises and conclusions are listed separately. It is not always easy to make such reconstructions, and various choices must be made. In principle, three different criteria must be taken into consideration:

- interpretation
- charity
- simplicity

### *Interpretation:*

Attempt to interpret the presentation in the original fragment as faithfully as possible (and, if applicable, the rest of the text or the broader debate, to check your interpretation).

### *Charity:*

Attempt to present as strong a version of the argument as possible. That is, describe the premises as plausibly as possible and add any missing steps.

Suppose someone says:

“Linda is not at home because she has class.”

If you want to evaluate this argument, you will also have to check two steps, namely, if Linda is in class and if the class is not at her home:

- (1) Linda has class.
- (2) If she has class, then Linda is there.
- (3) If Linda is in class, she is not at home.
- (4) Therefore: Linda is not at home.

(2) and (3) are called “hidden” or “implicit” premises.

In everyday contexts, it can often be assumed that these hidden premises are taken for granted and do not always need to be made explicit.

In philosophical contexts, things are typically more complicated, and then it is useful to make hidden premises explicit.

Example:

“I know that the world outside my consciousness exists because I know that this is a hand.”

In this case, it is useful to make the hidden premise explicit:

If I know that this is a hand, then I know that the world outside my consciousness exists.

### *Simplicity:*

Attempt to make as simple a version of the argument as possible. A reconstruction should make a text more transparent, not more complex. Focus on the most important steps and put sub-arguments aside if they make your reconstruction too complex.<sup>5</sup>

---

<sup>5</sup> A variant of “Ockham’s razor.”

Please note: the three criteria can conflict with each other, and then you will have to make choices. Arguments can often be reconstructed in several ways, any of which can yield a complete score on an exam *as long as you explain your choices and give reasons for them*.

The following obtains as a rule of thumb: in historical philosophy, the emphasis lies on interpretation (on understanding rather than revising texts), and in systematic philosophy on the principle of charity (on improving arguments and discussions, rather than only the interpretation of what is already present). But it would be wrong to think that the other criteria are then completely irrelevant.

The principle of charity is also important in debates. If you give a weak version of your opponent's argument and then critique it, you are then committing the "straw man fallacy." In that case, your critique fails for a stronger version of that argument.

//

Example:

Victor Lamme and Herman Philipse dispute the existence of free will. Lamme (neuroscientist) is opposed, and Philipse (philosopher) is a proponent.<sup>6</sup>

Reconstruction of Lamme's argument:

- (1) Our brain is influenced by factors we are not aware of.
- (2) Our brain determines our behaviour (and not the reasons we ourselves give for our behaviour).
- (3) If our brain is influenced by factors we are not aware of and if our brain determines our behaviour, then free will does not exist.
- (4) Therefore: free will does not exist.

Phlipse shows that (2) is ambiguous. Is it *always* the case that our brain determines our behaviour, or only *sometimes*?

If we work solely from the criterion of interpretation, we should perhaps choose the latter reading. After all, the cases that Lamme discusses show that our brain determines our behaviour *in those cases* and not *in all cases*.

But if (2) is weakened, (3) must then be weakened as well:  
"If our brain is influenced and if our brain *sometimes* determines our behaviour, then free will does not exist."

And this is not a strong premise. Why should free will not exist if our behaviour is determined only now and then by our brain (and now and then by our own reasons as well)?

Thus, based on the principle of charity, we should present Lamme's argument as follows:

- (1) Our brain is influenced by factors we are not aware of.

---

<sup>6</sup> Lamme's [video](#); Philipse's [video](#) (in Dutch, use subtitles).

- (2) Our brain *always* determines our behaviour (and never the reasons that we ourselves give for our behaviour).
- (3) If our brain is influenced by factors we are not aware of and if our brain *always* determines our behaviour, then free will does not exist.
- (4) Therefore: free will does not exist.

The disagreement between Lamme and Philipse can now be precisely identified: they disagree about (2).

Please note: (3) is a hidden premise in Lamme's argument, which is an assumption about free will that he does not make explicit ("free will exists only if our behaviour is sometimes influenced by the reasons that we ourselves give for our behaviour").<sup>7</sup>

//

Assignment 2.1:

Read sections 1-6 from:

Kagan, S. 2011. Do I Make a Difference? *Philosophy & Public Affairs* 39.

Reconstruct Kagan's argument in these sections.

Please note: simplicity is important in this assignment. Do not be distracted by details and only identify the most important steps that Kagan makes.

//

In this paper, Kagan explores a problem for consequentialism, which he defines as follows:

Individual act A by agent S is a morally condemnable if (and only if<sup>8</sup>):  
The consequences would have been better if S had not done A.

Example:

You pollute a stream from which a girl drinks and dies as a result. Your behaviour is morally condemnable because the girl dies, and if you had not polluted the stream, the girl would not have died.

According to Kagan, a problem arises in "collective action cases." In such cases, many people *together* cause some harm. But specific individuals make no difference to the outcome.

Kagan' central example is factory farming:

Whether or not you buy a chicken in the shop makes no difference with respect to production in factory farms. The number of chickens that are raised and slaughtered does not decrease if you do not decide to buy chickens any more. After all, there are enough other consumers who will buy them.

Kagan describes the problem for consequentialism in the following crucial passage:

---

<sup>7</sup> Philipse does indicate explicitly what he understands by free will (check for yourself).

<sup>8</sup> We will discuss the meaning of these terms later.



The problem, in effect, is this: consequentialism condemns my act only when my act makes a difference. But in the kind of cases we are imagining, my act makes no difference, and so cannot be condemned by consequentialism—even though it remains true that when enough such acts are performed the results are bad. Thus consequentialism fails to condemn my act. (2011: 108)

Stated in steps:

- (1) Consequentialism condemns individual act A only if A makes a difference.
- (2) A does not make a difference in collective action cases.
- (3) Therefore: consequentialism cannot condemn A in such cases (which constitutes a problem for consequentialism).

Please note: as a rule, you should not use all the sentences from a text for your reconstruction. Some sentences provide further explanation, or give details that are less important, and those do not have to appear in your reconstruction.

The structure of Kagan's paper, then, in terms of (1)-(3):

- Section 3: a defence of (2) on the basis of examples.
- Section 4: a defence of (1) against four objections (namely, that A could be condemned in a different consequentialist way, which Kagan disputes).
- Section 5: a defence of (2) against an objection (namely, that A does make a difference, but that difference is imperceptible).
- Section 6: the introduction of a different objection to (2) (namely, that A does make a difference if it is a "triggering case" and hits a threshold).

//

Types of arguments:

*Simple*

A conclusion is supported by one set of premises.

Example:

Linda is not at home, for she has class, and if she has class, she is not at home.

*Complex and compound*

Premises support a conclusion via intermediate steps (sub-conclusions).

Example:

Linda is not at home, for she has class, and she has class because the instructor says she does.

*Complex and multiple*

Different sets of premises support a conclusion.

Example:

Linda is not at home, for she has class and no one is opening the door.

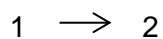
This differs from the simple argument in that there are now two different grounds for the conclusion.

The same typology obtains for the contra arguments (arguments *against* a standpoint instead of for it).

//

Arguments can be reproduced in an *arrow diagram*. In this kind of diagram, the numbers represent the different propositions (in the argument). And arrows pointing right ( $\rightarrow$ ) indicate support for a proposition (a pro argument).

Example:



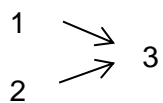
Here, 1 supports 2.

Example of a complex and compound argument:



Here, 1 supports 2 and 2 supports 3.

Example of a complex and multiple argument:



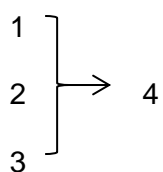
Here, 1 supports 3 and 2 supports 3.

Take Singer's argument:

- (1) If you can act against famine in the world by giving substantially, and you do not have to sacrifice something of comparable moral importance, then you must give substantially.
- (2) You can act against famine in the world by giving substantially.
- (3) You do not have to sacrifice anything of comparable moral importance.
- (4) Therefore: you have to give substantially (say, 10% of your income).

This is a simple argument: the conclusion is supported by one set of premises (1, 2, and 3 together form one pro argument 4).

The arrow diagram for the above argument:



It is not only pro arguments that can be presented in such diagrams; contra arguments can also be presented in the same way. In that case, the arrow pointing left ( $\leftarrow$ ) is used.

Suppose you do not agree with Singer's conclusion (4). You will then have to develop an argument against 1, 2, or 3.

Example:

3  $\leftarrow$  5

Here, 5 undermines 3.

In principle, you can reproduce complete debates (with all pro and contra arguments a certain standpoint) in an arrow diagram.

People are, generally speaking, visually oriented, and such visualisations can therefore help.

//

Assignment 2.2:

Determine what type of arguments are used in Kagan's text.

The main arguments are usually simple:

- Section 2: consequentialism condemns A only if A makes a difference, and A does not make any difference in collective action cases; therefore: consequentialism cannot condemn A in such cases.
- Section 4: consequentialism cannot condemn A in a different way; therefore, consequentialism condemns A only if A makes a difference. But the argument for this premise is complex and multiple: Kagan gives four different arguments why consequentialism cannot condemn A on different grounds.)
- Section 5: A makes perhaps an imperceptible difference in collection action situations, but if A makes an imperceptible difference, then that difference is morally irrelevant; therefore, A does not make any morally relevant difference in collective action cases.
- Section 6: A makes a difference if A is a triggering case; therefore, consequentialism can condemn A if it is a triggering case (or if you do not know if it is a triggering case).

Assignment 2.3:

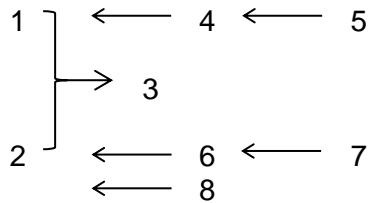
Make an arrow diagram that shows the most important steps of Kagan's argument.

Giving the propositions a number:

1. Consequentialism condemns A only if A makes a difference.
2. A does not make any difference in collective action cases.
3. Consequentialism cannot condemn A in collective action cases.
4. Consequentialism condemns A in other circumstances as well.
5. [Other arguments against 4.]
6. A does make a difference in collective action cases, namely, an imperceptible one.

7. [Other arguments against 6.]
8. A does make a difference in collective action cases, namely, if A is a triggering case.

Arrow diagram:



In this diagram, we can see what is undermined by what:

- 5, 7, and 8 are *not* undermined;
- 1 is not countered by 4 because 4 is itself undermined by 5;
- the attack on 2 is multiple (namely, from 6 and 8);
- 2 is not undermined by 6 because 6 is itself undermined by 7;
- 2 is countered by 8;
- thus, pro argument 3 is undermined by 8.

//

Usefulness:

If you master the technique of making reconstructions and corresponding arrow diagrams, you can understand a text *much* more quickly.

And then you can see if you agree with the argument.

Think what a difference it would make if a reconstruction of the argument would be available for every text that is assigned in some course.

As the editors of *Just the Arguments* put it:<sup>9</sup>

It was our experience that for almost any given philosophy class that we took as undergraduates, there were only a handful of arguments, totaling no more than a few pages of carefully crafted notes, that we needed to know. We imagined a rolodex of arguments in front of us, which we could spin through with ease to find the argument and move on. ... “Show me the argument” is the battle cry for philosophers. ... When things become serious, one wants *just the arguments*. (p. 1)

//

Assignment 2.4:

Choose a philosophical text that interests you and reconstruct the argument of a passage from that text and make a corresponding arrow diagram.

Do not make it too easy for yourself: choose a challenging passage with a high argumentative density.

---

<sup>9</sup> Bruce, M. & S. Barbone 2011. *Just the Arguments*. Wiley.

But do not make it too difficult either: choose a short passage of ~200 words so that you can analyse it thoroughly.

Give a brief explanation of the choices you made in your reconstruction (in terms of the criteria of interpretation, charity, simplicity).

Please note: if the author uses several sub-arguments, then you can focus on the main line in your reconstruction. After all, your reconstruction must make the text more transparent (not more complicated).

For inspiration, check out *Just the Arguments*. This book gives reconstructions of arguments from 100 well-known philosophical texts. Please note: if you choose one of these texts, you still have to make your *own* reconstruction. The reconstructions in this book are not always straightforward, and the point is that you *yourself* need to think about the three criteria (interpretation, charity, simplicity) and explain this process.

### 3. Reading

Some texts do not need too much context to be able to follow the argument (such as the texts in §1). But other texts are less accessible, and reading them can be a difficult task.

Reading has many aspects to it, and we will concentrate on two of them in what follows:

- framed reading
- unbiased reading

Framed reading is reading texts with a certain discussion in one's mind. The purpose of such framed reading is to place the text in that discussion. This is an important skill. It is quite possible in many philosophical texts to overlook relevant issues, or even to miss the whole argument in a text, if you are not aware of the wider discussion.

In principle, you can always ask for such a frame for all texts you have to read, that can help you read the text.

But framed reading also has a downside—namely, that you are biased about the text. You may expect the author to say this or that (because that is the mainstream reading), but the author is actually more nuanced.

In short, it is important not only to be able to read in a frame but also not to take that frame too strictly and to remain self-critical.

//

Read the following text:<sup>10</sup>

Morally speaking it may matter a great deal how a death comes about, whether from natural causes, or at the hands of another, for example. Does it matter whether a man was killed or only let die? A great many people think it does: they think that killing is worse than letting die. And they draw conclusions from this for abortion, euthanasia, and the distribution of scarce medical resources. ... And isn't what they mean by it confirmed by the fact—isn't it a fact?—that in the following case, Charles must not kill, that he must instead let die:

- (3) Charles is a great transplant surgeon. One of his patients needs a new heart, but is of a relatively rare blood-type. By chance, Charles learns of a healthy specimen with that very blood-type. Charles can take the healthy specimen's heart, killing him, and install it in his patient, saving him. Or he can refrain from taking the healthy specimen's heart, letting his patient die.

I should imagine that most people would agree that Charles must not choose to take out the one man's heart to save the other: he must let his patient die. And isn't what they mean by it further

---

<sup>10</sup> Thomson, J. J. 1976. Killing, Letting Die, and the Trolley Problem. *Monist* 59.

confirmed by the fact—isn't it a fact?—that in the following case, David must not kill, that he must instead let die:

- (4) David is a great transplant surgeon. Five of his patients need new parts—one needs a heart, the others need, respectively, liver, stomach, spleen, and spinal cord—but all are of the same, relatively rare, blood-type. By chance, David learns of a healthy specimen with that very blood-type. David can take the healthy specimen's parts, killing him, and install them in his patients, saving them. Or he can refrain from taking the healthy specimen's parts, letting his patients die.

If David may not even choose to cut up one where *five* will thereby be saved, surely what people who say "Killing is worse than letting die" mean by it must be right!

On the other hand, there is a lovely, nasty difficulty which confronts us at this point. Philippa Foot says—and seems right to say—that it is permissible for Edward, in the following case, to kill:

- (5) Edward is the driver of a trolley, whose brakes have just failed. On the track ahead of him are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Edward can turn the trolley onto it. Unfortunately there is one person on the right-hand track. Edward can turn the trolley, killing the one; or he can refrain from turning the trolley, killing the five.

If what people who say "Killing is worse than letting die" mean by it is true, how is it that Edward may choose to turn that trolley? Killing and letting die apart, in fact, it's a lovely, nasty difficulty: why is it that Edward may turn that trolley to save his five, but David may not cut up his healthy specimen to save his five? I like to call this the trolley problem, in honor of Mrs. Foot's example.

Mrs. Foot's own solution to the trolley problem is this. We must accept that our 'negative duties', such as the duty to refrain from killing, are more stringent than our 'positive duties', such as the duty to save lives. If David does nothing, he violates a positive duty to save five lives; if he cuts up the healthy specimen, he violates a negative duty to refrain from killing one. Now the negative duty to refrain from killing one is not merely more stringent than the positive duty to save one, it is more stringent even than the positive duty to save five. So of course Charles may not cut up his one to save one; and David may not cut up his one even to save five. But Edward's case is different. For if Edward 'does nothing', he doesn't do nothing; he kills the five on the track ahead, for he drives right into them with his trolley. Whichever Edward does, turn or not turn, he kills. There is, for Edward, then, not a conflict between a positive duty to save five and a negative duty to refrain from killing one; there is, for Edward, a conflict between a negative duty to refrain from killing five and a negative duty to refrain from killing one. But this is no real conflict: a negative duty to refrain from killing five is surely more stringent than a negative duty to refrain from killing one. So Edward may, indeed must, turn that trolley.

Now I am inclined to think that Mrs. Foot is mistaken about why Edward may turn his trolley, but David may not cut up his

healthy specimen. I say only that Edward “may” turn his trolley, and not that he must: my intuition tells me that it is not required that he turn it, but only that it is permissible for him to do so. But this isn’t important now: it is, at any rate, permissible for him to do so. Why? Compare (5) with

- (6) Frank is a passenger on a trolley whose driver has just shouted that the trolley’s brakes have failed, and who then died of the shock. On the track ahead are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Frank can turn the trolley onto it. Unfortunately there is one person on the right hand track. Frank can turn the trolley, killing the one; or he can refrain from turning the trolley, letting the five die.

If Frank turns his trolley, he plainly kills his one, just as if Edward turns his trolley, he kills his one: anyone who turns a trolley onto a man presumably kills him. Mrs. Foot thinks that if Edward does nothing, he kills his five, and I agree with this: if a driver of a trolley drives it full speed into five people, he kills them, even if he only drives it into them because his brakes have failed. But it seems to me that if Frank does nothing, he kills no one. He at worst lets the trolley kill the five; he does not himself kill them, but only lets them die.

But then by Mrs. Foot’s principles, the conflict for Frank is between the negative duty to refrain from killing one, and the positive duty to save five, just as it was for David. On her view, the former duty is the more stringent: its being more stringent was supposed to explain why David could not cut up his healthy specimen. So by her principles, Frank may no more turn that trolley than David may cut up his healthy specimen. Yet I take it that anyone who thinks Edward may turn his trolley will also think that Frank may turn his. Certainly the fact that Edward is driver, and Frank only passenger could not explain so large a difference.

So we stand in need, still, of a solution: why can Edward and Frank turn their trolleys, whereas David cannot cut up his healthy specimen? One’s intuitions are, I think, fairly sharp on these matters. Suppose, for a further example, that

- (7) George is on a footbridge over the trolley tracks. He knows trolleys, and can see that the one approaching the bridge is out of control. On the track back of the bridge there are five people; the banks are so steep that they will not be able to get off the track in time. George knows that the only way to stop an out-of-control trolley is to drop a very heavy weight into its path. But the only available, sufficiently heavy weight is a fat man, also watching the trolley from the footbridge. George can shove the fat man onto the track in the path of the trolley, killing the fat man; or he can refrain from doing this, letting the five die.

Presumably George may not shove the fat man into the path of the trolley; he must let the five die. Why may Edward and Frank turn their trolleys to save their fives, whereas George must let his five



die? George's shoving the fat man into the path of the trolley seems to be very like David's cutting up his healthy specimen. But what is the relevant likeness?" (1976: 204-8)

This passage is the introduction to Thomson's "Killing, Letting Die, and The Trolley Problem." This is an extremely influential paper, i.e. despite the fact that not many people have read beyond the passage quoted above (at least, there exists hardly any discussion of the rest of the paper).

The paper is often cited in the debate between consequentialists and Kantians. The general lines of the debate can be put as follows.

Consequentialism:

If you have to choose between two alternatives, choose the act with the best consequences for the greatest number of people.

Kantian ethics:

Choose the act whereby you view everyone as a goal in themselves and not just as a means to achieve your own goals.

Assignment 3.1:

Read the above passage once again with the discussion between consequentialists and Kantians in mind, and discuss how that helps to put Thomson's cases in perspective.

Can you see the relevant factors that were cited in the theories (e.g. whether people are used only as a means) in the cases?

And, do these theories perhaps overlook certain factors that seem to be important in Thomson's cases?

//

You can apply consequentialism in such a way that, according to that theory, you may *always* intervene to save the five (thus, that David the surgeon and George on the bridge can do that as well).

Kantian ethics may be applied in such a way that, according to it, you may *never* intervene (thus, including Frank the passenger) if you use one person as a mere means to save five others.

If so, case (6) would pose a problem for Kantian ethics (that people may never be killed to save others), whereas cases (4) and (7) pose a problem for consequentialism (that the number of survivors must always be maximised).

Even more, it could be thought that you nevertheless have to choose one of those theories to remain "consistent," and that you thus have to question Thomson's intuitions in some of her cases.

But that would be a misconception. Of course, there are always people who think that one must *always* maximise the number of survivors or that one may *never* kill to save others. But there are also people, like Thomson, who think that things are more complex.

//

Assignment 3.2:

Read Thomson's *whole* paper and attempt to trace, in as unbiased and accurate a way as possible, what Thomson sees as the answer to the "trolley problem."

This problem is this:

Why is Frank (6) allowed to kill one person to save five, while David (4) and George (7) may not? Is there any relevant difference?

//

General tip:

Do not hold back from looking up the complete and original sources *yourself* and checking the references. This will help stop you from adopting commonplaces that rest on misconceptions. (Moreover, original sources are usually more nuanced and more intriguing than secondary literature about them.)

#### 4. Validity

Compare:

- (1) If it's raining, then the streets are wet.
- (2) It's raining.
- (3) Therefore: the streets are wet.

- (1) If the streets are wet, then it's raining.
- (2) It's raining.
- (3) Therefore: the streets are wet.

Does (3) follow in both cases from (1) and (2)?

Note: there is a subtle but crucial difference in (1).

Whether conclusions follow (or inferences can be drawn) is a question of *validity*. An argument is valid if the conclusion follows from the premises (regardless of whether the premises are true or not).

Validity is not the same as *soundness*. An argument is sound if the conclusion follows from the premises (thus if the argument is valid) *and* the premises are true (or at least plausible).

Please note: never say that a standpoint is *valid* or an argument *true*: validity is a property that belongs to arguments, and truth a property that belongs to standpoints.

Discussion can concern not only truth but also validity. We often do not know if premises are indeed true (and would need more information to determine this), but we can determine if certain conclusions follow *if* certain premises are true (which we then leave aside).

Example:<sup>11</sup>

I think that the premise is false, that the fetus is not a person from the moment of conception. A newly fertilized ovum, a newly implanted clump of cells, is no more a person than an acorn is an oak tree. But I shall not discuss any of this. For it seems to me to be of great interest to ask what happens if, for the sake of argument, we allow the premise. How, precisely, are we supposed to get from there to the conclusion that abortion is morally impermissible?  
(Thomson 1971: 48)

Thomson here asks how (1) is to lead to (2) and leaves aside the question of whether (1) is indeed true:

- (1) The fetus is a person.
- (2) Therefore: abortion is not allowed.

//

Consider:

- (1) I'm in if I'm giving a class.

---

<sup>11</sup> Thomson, J. J. 1971. A Defense of Abortion. *Philosophy & Public Affairs* 1.

- (2) I'm not giving a class.
- (3) Therefore: I'm not in.

This is invalid: the conclusion does not follow from the premises.<sup>12</sup>

More specifically, there are situations conceivable whereby the premises are true, but the conclusion is still not true.

Suppose I am also in if I am having lunch. There can be various reasons why I am in. Premise (1) states only that giving a class is one of the reasons (and premise (2) that that reason does not apply). But the argument says nothing further about other reasons for being in.

Check 1:

Suppose that I am indeed out.

Does that affect the validity of the argument?

No, the argument can be invalid even if the conclusion is factually true.

Check 2:

Suppose that I am indeed giving a class.

Does that affect the validity of the argument?

No, the argument can be valid even if a premise is false.

Compare:

- (1) If I'm giving a class, then I'm in.
- (2) I'm not in.
- (3) Therefore: I'm not giving a class.

This argument is valid. There are no situations conceivable in which the premises are true but the conclusions false (check for yourself).

Assignment 4.1:

Construct an everyday argument that

- is valid but not sound;
- sound and valid;
- invalid but has true premises;
- invalid and not sound.

//

In this chapter we will focus on validity and leave soundness aside. The question is: what is the difference between valid and invalid arguments?

The difference has to do with their *form*. Some arguments take on a logically invalid form, others a logically valid form.

Logically valid argument forms:

- Modus ponens
- Modus tollens
- Disjunctive syllogism

---

<sup>12</sup> For the time being, we will concentrate on "deductive" validity, and in particular the following definition: an argument is deductively valid if there is no situation conceivable (or possible) in which the premises are true and the conclusion false.

- Dilemma
- Conditional proof
- Reductio ad absurdum

*Imprint this list on your mind.*

These are the most important argument forms that philosophers use. And they differ in subtle ways from their invalid counterparts.

//

The first and perhaps most well-known argument form:

**Modus ponens**

- |                  |           |
|------------------|-----------|
| (1) If p, then q | prem      |
| (2) p            | prem      |
| (3) Therefore: q | from 1, 2 |

An argument form consists of a set of lines whereby each line is indicated as a premise (prem) or an inference. In the latter case, we indicate from which preceding lines the inference is drawn.

The letters 'p' and 'q' in those lines represent propositions. Thus, if 'p' stands for the proposition that Socrates is a human being, and 'q' for the proposition that Socrates is mortal, we obtain the following argument:

- (1) If Socrates is a human being, then Socrates is mortal.
- (2) Socrates is a human being.
- (3) Therefore: Socrates is mortal.

All arguments in this form are valid: there are no conceivable cases in which the premises are true but the conclusion is not. If (1) and (2) are true in this case, then Socrates must be mortal. If he is not, then (1) and/or (2) must be false.

Example from Singer:

- (1) If you can act against famine in the world by giving substantially, and you do not have to sacrifice something of comparable moral importance, then you must give substantially.
- (2) You can act against famine in the world by giving substantially, and you do not have to sacrifice anything of comparable moral importance.
- (3) Therefore: you have to give substantially.

Versus:

**Affirmation of the consequent**

- |                  |           |
|------------------|-----------|
| (1) If q, then p | prem      |
| (2) p            | prem      |
| (3) Therefore: q | from 1, 2 |

This resembles a modus ponens, with the sole difference that the order of 'p' and 'q' in (1) is reversed.

All arguments of this type are *invalid*.

Example:

- (1) If it's raining, then the streets are wet.
- (2) The streets are wet.
- (3) Therefore: it's raining.

The conclusion does not follow. If the streets are wet if it is raining, and the streets are indeed wet, that still does not mean that it is raining (e.g. rather, the streets may be wet due to flooding).

//

### **Modus tollens**

- |                      |           |
|----------------------|-----------|
| (1) If p, then q     | prem      |
| (2) Not-q            | prem      |
| (3) Therefore: not-p | from 1, 2 |

An example from Kagan's paper:

- (1) Consequentialism condemns an individual act A only if A makes a difference.
- (2) A makes no difference in collective action cases.
- (3) Therefore: consequentialism cannot condemn A in such situations.

As stated in §2, this is not Kagan's final position.

Another example:

- (1) The ancient Greeks were morally responsible for slavery only if they could have known better.
- (2) They couldn't have known better.<sup>13</sup>
- (3) Therefore: the ancient Greeks were not morally responsible for slavery.

Versus:

### **Negation of the antecedent**

- |                      |           |
|----------------------|-----------|
| (1) If q, then p     | prem      |
| (2) Not-q            | prem      |
| (3) Therefore: not-p | from 1, 2 |

Example:

- (1) If it's raining, then the streets are wet.
- (2) It's not raining.
- (3) Therefore: the streets are dry.

This resembles the modus tollens, with the sole difference that the order of 'p' and 'q' in (1) is reversed.

Again, this is an invalid argument. The conclusion does not follow. If the streets are wet if it rains, and it is not raining, the streets can still be wet (think again of flooding).

//

---

<sup>13</sup> This is disputable, see end of §8.

Sometimes philosophers say, “one person’s modus ponens is another’s modus tollens.” What might this mean?

Take Singer’s argument:

- (1) If you can act against famine in the world by giving substantially, and you do not have to sacrifice something of comparable moral importance, then you must give substantially.
- (2) You can act against famine in the world by giving substantially, and you do not have to sacrifice anything of comparable moral importance.
- (3) Therefore: you have to give substantially.

This argument has a modus ponens form and is therefore valid. One could claim that the conclusion is absurd and difficult to accept as true. That means you deny (3) as a premise and draw a different conclusion (based on a modus tollens):

- (1) If you can act against famine in the world by giving substantially, and you do not have to sacrifice something of comparable moral importance, then you must give substantially.
- (2) You do *not* have to give substantially.
- (3) Therefore: it is not the case that you can act against famine by giving substantially or not the case that you thereby do not have to sacrifice anything of comparable moral importance.

Of course, you cannot simply transform one argument into a different argument. The premises will still have to be argued for. If you accept (1), you will have to show that your (2) is more plausible than Singer’s (2).

//

### Disjunctive syllogism

- |                  |           |
|------------------|-----------|
| (1) p or q       | prem      |
| (2) Not-p        | prem      |
| (3) Therefore: q | from 1, 2 |

Example:

- (1) The solution to problem X is either consequentialism or Kantian ethics.
- (2) Consequentialism is not a solution to X.
- (3) Therefore: Kantian ethics is the solution to X.

//

### Dilemma

- |                  |              |
|------------------|--------------|
| (1) p or q       | prem         |
| (2) If p, then r | prem         |
| (3) If q, then r | prem         |
| (4) Therefore: r | from 1, 2, 3 |

‘p’ and ‘q’ are also called the “horns” of the dilemma here.

Example:

- (1) Either consequentialism or Kantian ethics is correct.
- (2) If consequentialism is correct, then slavery is morally wrong.
- (3) If Kantian ethics is correct, then slavery is morally wrong.
- (4) Therefore: slavery is morally wrong.

If premise (1) is false, the dilemma is called “false.”

//

From this point on we will distinguish between premises (prem) and assumptions (ass).

Premise:

A proposition that is accepted as true/plausible in an argument.

Assumption:

A proposition that is not accepted as true/plausible but only temporarily accepted in order to deduce something from it.

Please note: “assumption” is being used here as a technical term (i.e. a term with a specific meaning that deviates from other meanings it might have). E.g. it is *not* a premise in an argument that is not supported by further argument.

Assumptions play a role in two different argument forms: conditional proof and reductio ad absurdum.

//

### Conditional proof

- |                             |                                   |
|-----------------------------|-----------------------------------|
| (1) p                       | ass                               |
| (2) q                       | from 1 (with additional premises) |
| (3) Therefore: if p, then q | from 1, 2                         |

Example (which we will work out more fully in §10):

- (1) Determinism is true. [ass]
- (2) I am not responsible for my deeds.
- (3) Therefore: if determinism is true, then I am not responsible for my deeds.

It is important that determinism is not accepted here as true but only considered in order to see what follows.

//

### Reductio ad absurdum

- |                      |                                   |
|----------------------|-----------------------------------|
| (1) p                | ass                               |
| (2) q and not-q      | from 1 (with additional premises) |
| (3) Therefore: not-p | from 1, 2                         |

This argument form is useful if you want to undermine your opponent’s position.



The reasoning is this. Temporarily assume your opponent's position  $p$ . Then you show that  $p$  (possibly with other premises that your opponent also accepts) yields a contradiction (absurdity). But if  $p$  yields a contradiction, then  $p$  cannot be true, which you conclude.

Thomson's argument:

- (1) Someone's right to life is always more important than someone else's right to decide what happens in and with her body. [ass]
- (2) Therefore: the violinist's right to life is more important than my right to decide what happens with my body.
- (3) (2) is false (which yields a contradiction).
- (4) Therefore: (1) is false.

Please note: (1) is not one of Thomson's premises, but an assumption she takes from her opponent, only in order to undermine it.

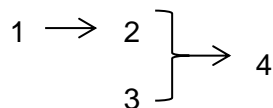
//

### Arrow diagrams

All logically valid argument forms are simple arguments (where the conclusion is supported by one set of premises), and can be simply reproduced in an arrow diagram.

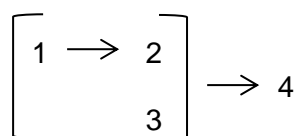
Matters are somewhat more complex if forms (pro or contra specific propositions) are combined.

But assumptions create problems. Thomson's argument above cannot be straightforwardly reproduced as follows:



Here, 1 seems to be part of an argument for 4, which cannot be the case because 4 is a negation of 1. The assumption 1 does not constitute any reason to accept the conclusion 4 (as a premise would). The same obtains for 2 and 3: they form a contradiction and do not seem to directly form a reason to accept 4.

What *does* constitute a reason to accept 4 is that 1 leads to a contradiction. We can indicate this by placing the argument from the assumption on between brackets:



//

### Paradoxes

Example:

- A is morally wrong.
- A is morally wrong only if A makes a difference.
- A does not make a difference.

A paradox is a set of propositions that seem to be plausible in themselves but cannot all be true.

The same applies here: the three propositions are plausible in the themselves (i.e. in the collective action cases discussed above), but taken together they are inconsistent, which means that at least one must be false.

A more complex example:<sup>14</sup>

- A debate can be settled.
- A debate can be settled only if there is a criterion accepted by both parties on which it can be settled.
- There is a criterion that is accepted by both parties only if the debate about the criterion can be settled.
- There can be an infinite regress of settling debates.

Please note: paradoxes themselves are not arguments. You can turn paradoxes into arguments by defending all individual propositions of the inconsistent set in question except for one.

Reductio ad absurdum plays an important role here:

- (1) A debate can be settled. [ass]
- (2) A debate can be settled only if there is a criterion accepted by both parties on which it can be settled.
- (3) There is a criterion that is accepted by both parties only if the debate about the criterion can be settled.
- (4) Therefore: an infinite regress of settling debates can be settled.
- (5) (4) is absurd (which yields a contradiction).
- (6) Therefore: no debate can be settled.

The regress pointed to in (4) is: first, there is a debate no. 1 on an arbitrary subject; then there is a debate no. 2 on the criterion on the basis of which debate no. 1 can be settled, and then there is a debate no. 3 on the criterion on the basis of which debate no. 2 must be settled, and so on.<sup>15</sup>

//

Assignment 4.2:

Construct an example of your own for each of the six logically valid argument forms discussed in the foregoing.

//

---

<sup>14</sup> Arguments like these are derived from the ancient sceptics; see Sextus Empiricus, *Outlines of Pyrrhonism*, 2.20. For more paradoxes, see Clarke, M. 2002. *Paradoxes From A to Z*. Routledge.

<sup>15</sup> Please note: none of the forms of argument we have seen bring us from (1)-(3) to (4).

Assignment 4.3:

Consider the following question posed by Plato, known as the “Euthyphro dilemma” (cf. *Euthyphro*, 10a):

“Is something moral because the gods approve of it, or do the gods approve of something because it is morally justified?”

Reconstruct the argument by means of the argument forms in this chapter (there are several possibilities).

What is immediately striking is that the standpoint is not made explicit in the quote from Plato. It is simply a question.

But assume that this is the conclusion (“not-p”):

It is not the case that something is morally justified just because the gods approve of it.

Which argument form leads us to this conclusion?

Option: reductio ad absurdum<sup>16</sup>

- |                      |                      |
|----------------------|----------------------|
| (1) p                | ass                  |
| (2) q                | from 1 <sup>17</sup> |
| (3) Not-q            | prem                 |
| (4) q and not-q      | from 2, 3            |
| (5) Therefore: not-p | from 1, 4            |

Key:

p: Something is morally justified just because the gods approve of it.

q: Slavery can be morally justified.

The idea behind this reconstruction:

If something is morally justified because the gods approve of it, then slavery can be morally justified. The gods could, after all, approve of slavery. But slavery cannot be morally justified, and therefore p yields a contradiction.

//

In short: how do you approach this?

1. Identify the author’s most important standpoint.
2. Look at what argument form you can use to arrive at that conclusion.
3. Identify the most important information in the text.
4. See if you can fill in the argument from step 2 with that information.

*This skill requires much practice!*

//

---

<sup>16</sup> For an alternative reconstruction see: *Just the Arguments*, p. 51.

<sup>17</sup> How does (2) follow from (1)? Hidden premise: ‘if p, then q’.

Forms of validity:

In what sense can a conclusion “follow” from premises?

Deductive:

An argument is *deductively* valid if the conclusion must be true if the premises are true (and otherwise deductively invalid).

There is in this case a necessary connection between the premises and the conclusion: there are no conceivable cases where the premises are true but the conclusion false.

This is the type of validity that we have discussed.

Some deductive arguments are also logically valid:

An argument is *logically* valid if it has a logically valid argument form (such as modus ponens and reduction ad absurdum).

An argument that is only deductively valid but not logically:

- (1) Socrates is a human being.
- (2) Therefore: Socrates is mortal.

The argument would be logically valid if we add an extra premise (so that it takes the form of a modus ponens):

- (1) Socrates is a human being.
- (2) If Socrates is a human being, then he is mortal.
- (3) Therefore: Socrates is mortal.

//

Inductive:

An argument is *inductively* valid if the conclusion is made plausible by the premises (and otherwise inductively invalid).

There is now no necessary connection (i.e. there are in principle cases conceivable where the conclusion is false but the premises true), but the premises do make the conclusion more probable.

Forms of induction:

- Generalisation
- Argument by analogy
- Abduction

Generalisation:

- |  |              |
|--|--------------|
| (1) a has property F                     | prem         |
| (2) b has property F                     | prem         |
| (3) c has property F                     | prem         |
| (4) Therefore: everything has property F | from 1, 2, 3 |

Please note: if you conclude too quickly that everything is F (without having carefully examined a representative collection), your conclusion will be called a “hasty generalisation”.

Argument by analogy:

- |                                      |           |
|--------------------------------------|-----------|
| (1) a has property F                 | prem      |
| (2) a and b share other properties   | prem      |
| (3) Therefore: b also has property F | from 1, 2 |

Example from Thomson:<sup>18</sup>

- (1) The violinist's right to life is not more important than my right to decide what happens in and with my body.
- (2) The situation of the violinist is comparable with the situation of abortion.
- (3) Therefore: the fetus' right to life is not more important than the right of the mother to decide what happens in and with her body.

Abduction:

- |                                     |           |
|-------------------------------------|-----------|
| (1) p                               | prem      |
| (2) q is the best explanation for p | prem      |
| (3) Therefore: q                    | from 1, 2 |

Examples:

- (1) I have two hands.
  - (2) That the external world exists is the best explanation for the fact that I have two hands.
  - (3) Therefore: the external world exists.
- 
- (1) The streets are wet.
  - (2) That it's raining is the best explanation for the fact that the streets are wet.
  - (3) Therefore: it's raining.

Please note: this argument looks much like the form "affirming the consequent":

- (1) The streets are wet.
- (2) If it's raining, then the streets are wet.
- (3) Therefore: it's raining.

Difference:

While the latter argument is deductively invalid (the conclusion does not have to be true even if the premises are true), the abductive argument is inductively valid (the conclusion is made plausible by the premises).

//

The difference between deductive and inductive arguments: *monotonicity*. Generally speaking, an argument is monotonic when premises can be added without making the argument invalid.

Inductively valid argument are not monotonic, for they can become invalid if a premise is added:

- (1) a has property F

---

<sup>18</sup> Another example: Singer's analogy between the child in the pond and poverty in the world (see §1).

- (2) b has property F
- (3) c has property F
- (4) d does not have property F [extra prem]
- (5) Therefore(?), everything has property F

- (1) a has property F
- (2) a and b share properties G and H
- (3) But: a and b do not share further properties [extra prem]
- (4) Therefore(?), b also has property F

In the latter case, there are relevant differences and that the analogy does not hold.

These arguments are no longer valid because the premises render the conclusion implausible.

This cannot happen in deductively valid arguments (which are monotonic).

## 5. Implications

Implications:

Propositions of the form “if [antecedent], then [consequent]”

Implications constitute the basis of the majority of arguments (think of modus ponens and modus tollens).

//

It is useful to remember that the following propositions mean the same (and are thus “equivalent”):

- if  $p$ , then  $q$
- $q$  if  $p$
- $p$  only if  $q$
- only if  $q$ ,  $p$

At first glance, this yields strange equivalencies:

- (i) If it's raining, then the streets are wet.
- (ii) It's raining only if the streets are wet.

Do (i) and (ii) mean the same? With (ii), one could ask: How is it a necessary condition for rain that the streets be wet? Does that not confuse cause (rain) and effect (wet streets)?

And what e.g. if the streets are protected by a covering? In that case (ii) is false. It can be raining, even if the streets stay dry because of the covering.

However, this does not imply that (i) and (ii) are not equivalent. After all, if there is a covering, then (i) is also false. Equivalency has to do with whether propositions are true in the same circumstances, and that holds for (i) and (ii).

Note: the above propositions are *not* equivalent with:

- If  $q$ , then  $p$
- $q$  only if  $p$

The direction of the implication is crucial. As we have seen, different things can follow from if “ $p$ , then  $q$ ” and “if  $q$ , then  $p$ .”

//

We easily make mistakes when it comes to implications.

Assignment 5.1:

There are four cards on the table:

- a card with the number 3
- a card with the number 8
- a red card
- a green card

Every card has a number and, on the other side, a colour.

Which card or cards do you have to turn over to see if the following implication is true?

“If a card has an even number, then the other side is red.”

This experiment is called “Wason Selection Task,” and only 4% do this test correctly.<sup>19</sup> The philosopher’s task is to be among those 4% and to be able to detect the errors in reasoning!

//

With the argument forms from §4 we have the means to solve Wason’s experiment.<sup>20</sup>

Card with 3:

This card does not need to be turned over. For, nothing follows from the premises “if even, then red” and “odd” (given the argument forms discussed).

Card with 8:

This even card does need to be turned over. Based on modus ponens, it follows that the other side has to be red (if even, then red; even; therefore red). This will need to be checked.

Red card:

This card does not need to be turned over. For, nothing follows from the premises “if even, then red” and “red” (given the argument forms discussed).

Green card:

This card must be turned over. It follows, based on modus tollens, that the other side must be odd (if even, then red; not-red; therefore odd). This needs to be checked.

//

Assignment 5.2:

Determine what mistake people make who choose the card with 3 or the red card.

Card with 3:

You know that “if even, then red” and “odd.” If you think that something interesting follows from this (in particular, “not red”), that is called negating the antecedent.

Red card:

You know that “if even, then red” and “red”. If you think that something interesting follows from this (in particular, “even”), that is called affirming the consequent.

---

<sup>19</sup> Wason, P. C. 1968. Reasoning About a Rule. *Quarterly Journal of Experimental Psychology* 20.

<sup>20</sup> Short answer: “if p, then q” is false only if p is true and q false. That’s why these two statements need to be checked.



Explain the experiment also in terms of:

- If  $p$ , then  $q$ .
- If someone drinks alcohol, then that person is over 18 years old.

//

Assignment 5.3:

Carry out Wason's experiment with someone who is not following this course and then attempt to explain the solution to that person. Briefly report your findings.

## 6. Propositional logic I

To check more precisely if and how arguments are logically valid, we can use formal systems. Here we will discuss such a system: propositional logic.

We will do this in two separate steps:

Step 1: translating sentences in everyday language into propositional logic.

Step 2: making deductions by means of propositional logic.

We will treat step 1 in this chapter, and step 2 in the next.

Propositional logic is a formal language in which propositions are symbolised by letters:  $p$ ,  $q$ ,  $r$ , etc.

The following symbols are also used:<sup>21</sup>

not/negation:	$\neg$
and/conjunction:	$\wedge$
or/disjunction:	$\vee$
if then/implication:	$\rightarrow$
equivalence:	$\leftrightarrow$
therefore/inference:	$/$

Let us use the following translation key:

$p$ : It's raining.

$q$ : The weather gods are angry.

Example of translations:

It's not raining.

$\neg p$

It's raining and the weather gods are angry.

$p \wedge q$

It's raining or it's not raining.

$p \vee \neg p$

If it's raining, then the weather gods are angry.

$p \rightarrow q$

If the weather gods are angry, then it's raining.

$q \rightarrow p$

The conjunction of the last two sentences can also be stated using the equivalence symbol:<sup>22</sup>

$p \leftrightarrow q$

Finally, we use parentheses to avoid ambiguities. For example:

---

<sup>21</sup> Note that alternative symbols are used too. To insert symbols in Word: Insert > Symbol > More symbols > Font "Symbol."

<sup>22</sup> This sign is used for definitions where the left side is necessary and sufficient for the right side, and conversely; see §12.

$\neg p \rightarrow q$   
 $\neg(p \rightarrow q)$

Translation:

If it doesn't rain, then the weather gods are angry.

It is not the case that if it rains, the weather gods are angry.

These sentences say different things:

The first says: if it is dry, we have to do with angry gods.

The second says: it can be the case that it's raining and that the gods are happy.

If you think that a statement in everyday language is ambiguous: give both readings and both translations.

//

Assignment 6.1:

Translate the following sentences into propositional logic:

- a. Zeus is a weather god and Aphrodite not.
- b. Zeus and Aphrodite are Greek gods.
- c. It will rain tomorrow or the next day.
- d. If it doesn't rain this week or the next, then the weather gods are sad.
- e. If it doesn't rain and the weather gods are sad, then unemployment will rise.
- f. The economy is improving but unemployment is not decreasing.
- g. If Zeus or Poseidon do not win, then Aphrodite wins.
- h. If Aphrodite wins, then Zeus and Poseidon do not.

Some instructions:

Make a translation key explicit. Without a key, your translation will mean nothing.

Check if you have translated all negations (and do not take a key where e.g. "p" stands for "Aphrodite is not a weather god").

Check the directions of any arrows.

Check if you have not confused conjunctions and disjunctions (in e.g. d and g).

Compare:

- (i) It will not rain this week or next week.
- (ii) It will not rain this week and next week.

The translations:

- (i)  $\neg p \vee \neg q$  or  $\neg(p \wedge q)$
- (ii)  $\neg p \wedge \neg q$  or  $\neg(p \vee q)$

Where:

p: It will rain this week.

q: It will rain next week.

(i) is applicable in d. The weather gods are sad if it does not rain at all.

Which is applicable in g? It is plausible to assume that only one god can win. This means that both Zeus and Poseidon have to lose, and then we choose (ii). If both lose, Aphrodite wins, but not if only Zeus or only Poseidon loses.

Again: if you think a statement in everyday language is ambiguous, explain your reading and give the appropriate translation.

//

Assignment 6.2:

Translate the following sentences:

- a. Unemployment can only be decreased if the economy improves.
- b. We will go outside if and only if the weather gods have passed by.
- c. The weather gods are happy if the economy improves.
- d. Unless the weather gods strike, it will rain tomorrow.

//

Assignment 6.3:

Translate the following fragment by Thomson into propositional logic.

Note: ignore sentences in your translation that are irrelevant to arriving at the conclusion. As always, find the latter first.

I propose, then, that we grant that the fetus is a person from the moment of conception. How does the argument go from here? Something like this, I take it. Every person has a right to life. So the fetus has a right to life. No doubt the mother has a right to decide what shall happen in and to her body; everyone would grant that. But surely a person's right to life is stronger and more stringent than the mother's right to decide what happens in and to her body, and so outweighs it. So the fetus may not be killed; an abortion may not be performed.

It sounds plausible. But now let me ask you to imagine this. You wake up in the morning and find yourself back to back in bed with an unconscious violinist. A famous unconscious violinist. He has been found to have a fatal kidney ailment, and the Society of Music Lovers has canvassed all the available medical records and found that you alone have the right blood type to help. They have therefore kidnapped you, and last night the violinist's circulatory system was plugged into yours, so that your kidneys can be used to extract poisons from his blood as well as your own. The director of the hospital now tells you, "Look, we're sorry the Society of Music Lovers did this to you—we would never have permitted it if we had known. But still, they did it, and the violinist now is plugged into you. To unplug you would be to kill him. But never mind, it's only for nine months. By then he will have recovered from his ailment, and can safely be unplugged from you." Is it morally incumbent on you to accede to this situation? No doubt it would be very nice of you if you did, a great kindness. But do you have to accede to it? What if it were not nine months, but nine years? Or longer still? What if the

director of the hospital says, "Tough luck, I agree, but you've now got to stay in bed, with the violinist plugged into you, for the rest of your life. Because remember this. All persons have a right to life, and violinists are persons. Granted you have a right to decide what happens in and to your body, but a person's right to life outweighs your right to decide what happens in and to your body. So you cannot ever be unplugged from him." I imagine you would regard this as outrageous, which suggests that something really is wrong with that plausible-sounding argument I mentioned a moment ago. (1971: 48-9)

## 7. Propositional logic II

Now that we can translate sentences into propositional logic, we can continue on to the following step: making deductions.

We will use a system of 10 rules for this.<sup>23</sup>

Please note: the system is not difficult in itself, but it takes time to master it. Practice a great deal, and don't hesitate to ask for help.

The first two rules for conjunction:

Conjunction introduction ( $I_{\wedge}$ ):<sup>24</sup>

1.  $\alpha$
2.  $\beta$
3.  $\alpha \wedge \beta$        $I_{\wedge} (1, 2)$

Conjunction elimination ( $E_{\wedge}$ ):

1.  $\alpha \wedge \beta$
2.  $\alpha$ <sup>25</sup>       $E_{\wedge} (1)$

These rules are self-explanatory. If you know that I exist and you exist, then you know that we both exist. And conversely, if you know that we both exist, you know that you exist (and that I exist).

Implication introduction ( $I_{\rightarrow}$ ):

1.  $\alpha$       *ass*
2.  $\beta$
3.  $\alpha \rightarrow \beta$        $I_{\rightarrow} (1, 2)$

We know this already from §4: conditional proof.

Implication elimination ( $E_{\rightarrow}$ ):

1.  $\alpha \rightarrow \beta$
2.  $\alpha$
3.  $\beta$        $E_{\rightarrow} (1, 2)$

We know this already: modus ponens.

Disjunction introduction ( $I_{\vee}$ ):

1.  $\alpha$
2.  $\alpha \vee \beta$        $I_{\vee} (1)$

Example: if you know that it is raining, then you also know that it is raining or that something else is true (e.g. that soup is nice).

---

<sup>23</sup> This system ("natural deduction") and the exercises are taken from: Wim de Jong 2010. *Argumentatie en formele structuur. Basisboek logica*, chapter 5. Boom. We use natural deduction (rather than truth tables) as it is closer to actual philosophical reasoning.

<sup>24</sup> In these rules, the ' $\alpha$ ' and ' $\beta$ ' can be filled in for simple propositions (such as ' $p$ ') but also for more complex propositions (such as " $p \wedge q$ ").

<sup>25</sup> The same goes for  $\beta$ .

We are not familiar with this one yet, and it is not quite so intuitive. But the idea is that the disjunction can be deduced because it is true regardless of *which* disjunct is true.

Disjunction elimination ( $E\vee$ ):

1.  $\alpha \vee \beta$
2.  $\alpha \rightarrow \gamma$
3.  $\beta \rightarrow \gamma$
4.  $\gamma$   $E\vee (1, 2, 3)$

We know this as: dilemma.

Negation introduction ( $I\neg$ ):

1.  $\alpha$   $ass$
2.  $\beta \wedge \neg\beta$
3.  $\neg\alpha$   $I\neg (1, 2)$

We know this as: reductio ad absurdum.

Negation elimination ( $E\neg$ ):

1.  $\alpha$
2.  $\neg\alpha$
3.  $\beta$   $E\neg (1, 2)$

If you know that it is raining and that it is *not* raining, you can conclude anything.

This rule  $E\neg$  does not feel immediately intuitive:

How, from the proposition that it is raining and the proposition that it is not raining, can I deduce that e.g. soup is nice (or any other conclusion)?

Remember, however, when arguments are deductively valid: if there is no conceivable situation where the premises are true but the conclusion is not. In this case, there is no situation conceivable where both premises are true ( $\alpha$  and  $\neg\alpha$ ), and so the argument is valid.

Note the difference with  $I\neg$ :

$E\neg$  eliminates the negation (rather than introduces one).

In philosophy,  $I\neg$  will be encountered more often than  $E\neg$  (and  $I\neg$  is intuitive).

Double negation elimination (elim):

1.  $\neg\neg\alpha$
2.  $\alpha$   $elim (1)$

If you know that it is not the case that it is not raining, then you know that it is raining.

Too many negations make it difficult (or even impossible) to follow an argument, and you can remove them with this rule.

Repetition (rep):

1.  $\alpha$
2.  $\alpha$   $rep (1)$

If you know that it is raining, you can use this information later again in another argument, and you can do that via this rule.

//

In short, we already know some rules from §4:

- $E\rightarrow$ : modus ponens
- $E\vee$ : dilemma
- $I\rightarrow$ : conditional proof
- $I\neg$ : reductio ad absurdum

The alert reader will see that two are missing: modus tollens and disjunctive syllogism. These can, however, be deduced from other rules.

Take modus tollens. Translated, it reads as follows:  $p \rightarrow q; \neg q / \neg p$ . The deduction, then, is this:

- |    |                   |                       |
|----|-------------------|-----------------------|
| 1. | $p \rightarrow q$ | prem                  |
| 2. | $\neg q$          | prem                  |
| 3. | $p$               | ass                   |
| 4. | $q$               | $E\rightarrow (1, 3)$ |
| 5. | $q \wedge \neg q$ | $I\wedge (2, 4)$      |
| 6. | $\neg p$          | $I\neg (3, 5)$        |

How do you make such deductions yourself?

There is a fixed sequence of steps to follow:

Step 1:

Write down your premises (if you have them) and put the conclusion next to them so that you know where you want to go.

- |    |                   |      |                |
|----|-------------------|------|----------------|
| 1. | $p \rightarrow q$ | prem | Goal: $\neg p$ |
| 2. | $\neg q$          | prem |                |
| 3. | ...               |      |                |

Step 2:

See if you can arrive at the conclusion via the 10 rules.

That is not the case here. We cannot e.g. apply  $E\rightarrow$  to 1 and 2.

Step 3:

What form does the conclusion have: is it a negation, an implication, etc.?

(If the conclusion has several symbols, choose the main symbol. E.g. the main symbol of " $\neg(p \wedge q)$ " is the negation, not the conjunction.)

There is a separate instruction for every form:

<i>Goal</i>	<i>Instruction</i>
$\neg\alpha$	Take $\alpha$ as an assumption and attempt to deduce a contradiction (with the premises you have), and close with $I\neg$ .



$\alpha$	Take $\neg\alpha$ as an assumption and try to deduce a contradiction (with the premises you have), and close with $I_{\neg}$ and elim.
$\alpha \rightarrow \beta$	Take $\alpha$ as an assumption and attempt to deduce $\beta$ from it (with the premises you have) and close with $I_{\rightarrow}$ .
$\alpha \vee \beta$	Attempt to deduce one of the two (from premises you have) and add the other with $I_{\vee}$ .
$\alpha \wedge \beta$	Attempt to deduce both (from premises you have) and combine them with $I_{\wedge}$ .

In our case, we have to do with a negation, thus we will continue as follows:

- |                      |      |                         |
|----------------------|------|-------------------------|
| 1. $p \rightarrow q$ | prem | Goal: $\neg p$          |
| 2. $\neg q$          | prem |                         |
| 3. $p$               | ass  | Sub-goal: contradiction |

Now, according to the instruction, we need to look for a contradiction (i.e. " $\alpha \wedge \neg\alpha$ "). That is not far away: from 1 and 3 we get  $q$  (with  $E_{\rightarrow}$ ), in 2 we already have  $\neg q$ , and then we combine them together (with  $I_{\wedge}$ ):

- |                      |                          |
|----------------------|--------------------------|
| 4. $q$               | $E_{\rightarrow} (1, 3)$ |
| 5. $q \wedge \neg q$ | $I_{\wedge} (2, 4)$      |

We now have everything we need to deduce the conclusion:

- |             |                   |
|-------------|-------------------|
| 6. $\neg p$ | $I_{\neg} (3, 5)$ |
|-------------|-------------------|

We also say here that with  $I_{\neg}$  we have "closed" the assumption in 3 (more on this below).

*You need all the steps.* Don't think some steps are so obvious that you don't need to write them down. Propositional logic is like a computer that freezes up if a rule needs to be applied to something that isn't there.

//

#### Assignment 7.1:

Show that the following arguments are valid using propositional logic.

- $p \wedge q, p \rightarrow ((q \wedge p) \rightarrow r) / r$
- $p, p \rightarrow (p \rightarrow (p \rightarrow q)) / q$
- $p, p \rightarrow q, p \rightarrow r / q \wedge (r \wedge p)$
- $p \wedge q, p \rightarrow r / q \wedge r$
- $p \wedge q, p \rightarrow t, q \rightarrow s / t \wedge s$
- $p / q \rightarrow (p \wedge q)$
- $p \rightarrow q, q \rightarrow r / p \rightarrow r$
- $p \rightarrow r, q \rightarrow r / (p \wedge q) \rightarrow r$

7.1.a. as illustration:

To begin with, write down everything you have:

- |   |      |           |
|---|------|-----------|
| 1. $p \wedge q$                                 | prem | Goal: $r$ |
| 2. $p \rightarrow ((q \wedge p) \rightarrow r)$ | prem |           |

How do we arrive at the conclusion? There is quite an amount of information in the premises, so let us look at how far that will take us. In premise 2, we see “r” at the end of two implications. If we constantly find the antecedent, we will get (with E→) to where we want to be in the end. Let us first take “p” out of 1, then we can look at the consequent in 2:

- 3. p                                    E∧ (1)
- 4. (q ∧ p) → r                    E→ (2, 3)

We can then apply the same step: find the antecedent and then deduce the consequent:

- 5. q                                    E∧ (1)
- 6. q ∧ p                            I∧ (3, 5)
- 7. r                                    E→ (4, 6)

//

You are only allowed to introduce assumptions for the introduction of an implication (I→) and for the introduction of a negation (I¬).

All assumptions that you make along the way must be closed at a given moment (i.e. by one of these two rules).

Sometimes we will have to make several assumptions. In that case, it is crucial that we close them in the right way.

Example:

- 1. p → r                            prem                    Goal: p → (q → r)

Our goal is an implication, so we will assume the antecedent (of the “main implication”):

- 2. p                                    ass                    Sub-goal: (q → r)

Our sub-goal is again an implication, so we will therefore insert a second assumption.

- 3. q                                    ass                    Sub-goal: r

Our final sub-goal can now be found:

- 4. r                                    E→ (1, 2)

We will now close the assumptions, via I→ in this case. We close the final assumption first and then the assumption before that:

- 5. q → r                            I→ (3, 4)
- 6. p → (q → r)                    I→ (2, 5)

And so we have reached our goal.

Two rules for closing assumptions:

One: you can only close the last assumption made. In the above case, you first have to close the assumption in 3 before you can close the assumption in 2.

Two: you may no longer use assumptions if they have been closed. In the above case, the assumption in 3 may no longer be used if it has been closed via  $I\rightarrow$  in 5.

//

The argument form from §4 that we still have to look at is disjunctive syllogism. This can also be deduced but requires several steps, as well as several assumptions:

- |    |            |      |           |
|----|------------|------|-----------|
| 1. | $p \vee q$ | prem | Goal: $q$ |
| 2. | $\neg p$   | prem |           |

We see a disjunction in 1. If we can show that  $q$  is implied by both disjunctions, then we can conclude  $q$  with  $E\vee$ . In other words, we are going to look for  $p \rightarrow q$  and  $q \rightarrow q$ . Let us take  $p$  first:

- |    |     |     |               |
|----|-----|-----|---------------|
| 3. | $p$ | ass | Sub-goal: $q$ |
|----|-----|-----|---------------|

We now want to go to  $q$ . How do we do that? You can assume  $\neg q$  and deduce a contradiction.<sup>26</sup>

- |    |                   |                       |  |
|----|-------------------|-----------------------|--|
| 4. | $\neg q$          | ass                   |  |
| 5. | $p \wedge \neg p$ | $I\wedge$ (2, 3)      |  |
| 6. | $\neg\neg q$      | $I\neg$ (4, 5)        |  |
| 7. | $q$               | elim (6)              |  |
| 8. | $p \rightarrow q$ | $I\rightarrow$ (3, 7) |  |

This is the first part. The second is easier because we can immediately deduce  $p$  from  $q$ :

- |     |                   |                       |               |
|-----|-------------------|-----------------------|---------------|
| 9.  | $q$               | ass                   | Sub-goal: $q$ |
| 10. | $q$               | rep (9)               |               |
| 11. | $q \rightarrow q$ | $I\rightarrow$ (9, 7) |               |

Now we have everything we need to use the disjunction in 1 (note the notation):

- |     |     |                    |  |
|-----|-----|--------------------|--|
| 12. | $q$ | $E\vee$ (1, 8, 11) |  |
|-----|-----|--------------------|--|

All the steps together:

- |    |            |      |
|----|------------|------|
| 1. | $p \vee q$ | prem |
| 2. | $\neg p$   | prem |
| 3. | $p$        | ass  |
| 4. | $\neg q$   | ass  |
- 

<sup>26</sup> A shorter route: apply  $E\neg$  to 2 and 3, and derive  $q$  immediately from them.

- |     |                   |                          |
|-----|-------------------|--------------------------|
| 5.  | $p \wedge \neg p$ | $I_{\wedge}$ (2, 3)      |
| 6.  | $\neg\neg q$      | $I_{\neg}$ (4, 5)        |
| 7.  | $q$               | elim (6)                 |
| 8.  | $p \rightarrow q$ | $I_{\rightarrow}$ (3, 7) |
| 9.  | $q$               | ass                      |
| 10. | $q$               | rep (9)                  |
| 11. | $q \rightarrow q$ | $I_{\rightarrow}$ (9, 7) |
| 12. | $q$               | $E_{\vee}$ (1, 8, 11)    |

//

### Assignment 7.2:

- $\neg q / p \rightarrow (q \rightarrow r)$
- $p \rightarrow (q \vee \neg r) / (p \wedge r) \rightarrow q$
- $p \wedge q / \neg(\neg p \vee \neg q)$
- $p \rightarrow (q \vee r), \neg q \wedge \neg r / \neg p$

### 7.2.c. by way of illustration

We cannot move from premise " $p \wedge q$ " to conclusion " $\neg(\neg p \vee \neg q)$ ". How do we begin?

The conclusion does have a negation as its main symbol. The instruction then says: assume " $\neg p \vee \neg q$ " and attempt to deduce a contradiction (so that you can reach the conclusion with  $I_{\neg}$ ).

- |    |                      |      |                                  |
|----|----------------------|------|----------------------------------|
| 1. | $p \wedge q$         | prem | Goal: $\neg(\neg p \vee \neg q)$ |
| 2. | $\neg p \vee \neg q$ | ass  | Sub-goal: contradiction          |

What now? There is not much we can do with this premise at this time. But we do see a disjunction in the assumption. This can be used with the rule  $E_{\vee}$ :

$$\begin{array}{l} \neg p \vee \neg q \\ \neg p \rightarrow \dots \\ \neg q \rightarrow \dots \\ / \dots \end{array}$$

The question now is how do we fill " $\dots$ " in. As stated, we want to move towards a contradiction. We see " $p \wedge q$ " in the premise. If we can find  $\neg p$  or  $\neg q$ , we have a contradiction. Let us choose  $\neg p$ , we will then look for these:

$$\begin{array}{l} \neg p \rightarrow \neg p \\ \neg q \rightarrow \neg p \end{array}$$

We find this by always assuming the antecedent and attempting to deduce the consequent. That can be done immediately with the first implication:

- |    |                             |                          |                    |
|----|-----------------------------|--------------------------|--------------------|
| 3. | $\neg p$                    | ass                      | Sub-goal: $\neg p$ |
| 4. | $\neg p$                    | rep                      |                    |
| 5. | $\neg p \rightarrow \neg p$ | $I_{\rightarrow}$ (3, 4) |                    |

With the second implication, we will need an extra step. How do we proceed from  $\neg q$  to  $\neg p$ ? We will look again at what we have: we see “q” in the premise. If we take that out of 1, then we can move to  $\neg p$  via  $E\neg$ :

- |                                |                       |                    |
|--------------------------------|-----------------------|--------------------|
| 6. $\neg q$                    | ass                   | Sub-goal: $\neg p$ |
| 7. q                           | $E\wedge$ (1)         |                    |
| 8. $\neg p$                    | $E\neg$ (6, 7)        |                    |
| 9. $\neg q \rightarrow \neg p$ | $I\rightarrow$ (6, 8) |                    |

We now have everything we need to apply  $E\vee$ :

- |              |                   |
|--------------|-------------------|
| 10. $\neg p$ | $E\vee$ (2, 5, 9) |
|--------------|-------------------|

If we now take “p” out of 1 as well and bring the two together, we have our contradiction:

- |                       |                    |
|-----------------------|--------------------|
| 11. p                 | $E\wedge$ (1)      |
| 12. $p \wedge \neg p$ | $I\wedge$ (10, 11) |

Now we only need to close the assumption in 2 with  $I\neg$ , and we have arrived at our conclusion:

- |                                |                 |
|--------------------------------|-----------------|
| 13. $\neg(\neg p \vee \neg q)$ | $I\neg$ (2, 12) |
|--------------------------------|-----------------|

//

Assignment 7.3:

- $\neg(p \wedge q) / p \rightarrow \neg q$
- $(p \rightarrow q) \rightarrow p, p \rightarrow (p \rightarrow q) / q$
- $p \rightarrow (q \rightarrow r) / q \rightarrow (p \rightarrow r)$
- $/ (p \rightarrow \neg q) \rightarrow \neg(p \wedge q)$
- $\neg p \rightarrow q, (p \vee r) \rightarrow s, \neg q / s$
- $(p \wedge q) \rightarrow (r \rightarrow s), \neg s \wedge r / p \rightarrow \neg q$
- $\neg p \vee \neg q, r \rightarrow p, s \rightarrow q / \neg r \vee \neg s$

//

Assignment 7.4:

Show that an argument from a philosophical text of your choice is valid via propositional logic (use at least 4 different rules).

//

You have now learned something that you had perhaps not expected in your study of philosophy. Why, actually, should you be able to make these deductions?

We said in §4 that the quality of arguments partly depends on their validity, and we saw the most important deductive argument forms (modus ponens, etc.). We have just focused on a prominent formal system in which these argument forms are valid: propositional logic.

There is more to say about propositional logic, and there are many further formal systems, but we will not deal with them in this textbook.

If, in what follows, you see someone drawing a conclusion and it is not immediately clear whether that conclusion follows, use the tools that you have at your disposal: make a translation in terms of “p” and “q” and attempt to deduce the conclusion via the rules.

## 8. Soundness

If an argument is valid, it may still not be a good or sound argument. After all, the conclusion may follow from the premises, but the premises can still be implausible. We will now look at this other side of evaluating arguments.

//

Assignment 8.1:

Make a reconstruction of the argument in:

Rosen, G. 2004. Skepticism About Moral Responsibility. *Philosophical Perspectives* 18. [sections 4-6, pp. 298-304]

Context:

A classic form of scepticism concerning moral responsibility maintains that all our behaviour is determined by external causes, that we therefore have no possibility of acting otherwise, and that we therefore are not morally responsible for what we do.

Rosen's scepticism concerning moral responsibility is of another kind and not dependent on deterministic assumptions. His scepticism does not concern the excuse of compulsion ("I couldn't do anything else") but the one of ignorance ("I didn't know better"). According to Rosen, we are almost *always* excused on the basis of ignorance (an extreme conclusion!).

Reconstruction 1:

- (1) S is culpable for act A done out of ignorance only if S is culpable for her<sup>27</sup> ignorance.
- (2) S is culpable for her ignorance only if her ignorance is the result of a prior culpable act.
- (3) Her ignorance is not a consequence of a prior culpable act.
- (4) Therefore: S is not culpable for A.

Steps (1) and (2) are present almost literally in Rosen's paper, but not (3) and (4). Although they should follow naturally from the argument, Rosen's conclusion is somewhat weaker. Rosen states earlier that we *cannot establish with certainty* that S is culpable for previous behaviour, and therefore we cannot establish with certainty that S is culpable for A.

Reconstruction 2:

- (1) S is culpable for act A done out of ignorance only if S is culpable for her own ignorance.
- (2) S is culpable for her own ignorance only if her ignorance is a consequence of a prior culpable act.
- (3\*) We cannot establish with certainty that her ignorance is the result of a prior act by her.
- (4\*) Therefore: we cannot establish with certainty that S is culpable for A.

---

<sup>27</sup> Whether or not you agree with it, word use is accompanied by certain stereotypes. To prevent the use of sexist language, philosophers often use female pronouns.

The argument form cannot be ascertained immediately. It resembles a kind of modus tollens, but not completely.

First of all, modus tollens makes use of one implication, not two. Premises (1) and (2) can, however be taken as one:<sup>28</sup>

(1\*) S is culpable for act A done out of ignorance only if her ignorance is the result of a prior culpable act.

But that still does not get us from (1\*) to (4\*), for (4\*) concerns whether we can establish something with certainty and (1\*) says nothing about that. We need to modify this:

(1\*\*) We can establish with certainty that S is culpable for act A done out of ignorance only if we can establish with certainty that her ignorance stems from a prior culpable act.

(4\*) now follows from (3\*) and (1\*\*) via modus tollens. (4\*) is Rosen's sceptical conclusion.

//

Usefulness:

If you have identified all steps of the argument and checked them for validity, you can then see what options you have.

After all, if the argument is valid, the conclusion is true/plausible if the premises are true/plausible.

This means:

If you want to defend the conclusion, you have to defend all the premises.

And: if you do *not* want to accept the conclusion, you have to attack at least one of the premises.

//

Assignment 8.2:

See what we can introduce against Rosen's argument. Where might the argument be problematic?

Explain your answer by means of an example (of your own or an example from Rosen).

//

For convenience's sake, let's take reconstruction 1. This reconstruction has three premises: (1), (2), (3\*). Can we call at least one of them into question?

Premise (1) is plausible. Suppose you prepare an exotic dinner for your friends, and one of them turns out to be allergic to one of the herbs you

---

<sup>28</sup> After all, if X is necessary for Y, and Y necessary for Z, then X is also necessary for Z. We will look more closely at necessary conditions in §11.



have used in the food, and she is confined to her home for weeks because of the illness that results. Should you bear some of the blame? Let us assume that you knew nothing of the allergy, that your friends also knew nothing, and that you cannot be held responsible for your ignorance. If that is the case, then it seems you are not responsible for the allergic reaction either. If it is not your fault that you are ignorant, it is not your fault that you do certain things out of ignorance. Or: (1).

(2) is also plausible. You bear responsibility for your ignorance insofar as you are responsible for the fact that you did not make the effort to be better informed. If, for example, you could not be better informed, then you are also not responsible if you did not know better.

That leaves (3\*). Can we indeed not establish with certainty if one is culpable for a previous act?

Here is an example from Rosen:

Suppose a surgeon orders that her type A patient be transfused with type B blood, that she does this only because she is mistaken about the patient's blood type, and that she is mistaken about the blood type only because she neglected to double-check his chart (which had just been updated) immediately prior to surgery, as standard practice requires. Then we know this much. The surgeon is culpable for the bad transfusion (and the ensuing harm) only if she is culpable for her ignorance as to the patient's blood type, and she is culpable for her ignorance only if she is culpable for her negligent failure to double-check the chart. Now focus on this negligent failure, which is of course the failure to comply with a [duty of inquiry]. It is plausible that in this case, as in most similar cases, the negligent failure will itself be an act done from ignorance. When the time is ripe for the precaution, the agent will fail to take it only because she does not then think—and so does not then know—that she ought to be taking it. But if the failure is thoughtless in this sense then our principles entail that the agent will be culpable for it only if she is culpable for the ignorance that underlies it. Our principles then further entail that she will be culpable for this bit of ignorance only if she is culpable for omitting some required precaution to prevent it—e.g. if she is forgetful, asking one of her colleagues to remind her to check the chart. And here the same series of questions will arise again. ... For now the point is simply to convey how complex and arcane the inquiry can be when we set out to determine, in light of our principles, whether an agent is responsible for an action done from ignorance. (2004: 303)

In terms of a previously formulated reconstruction:

- (1) The surgeon is culpable for the wrong blood transfusion only if she is culpable for her ignorance concerning the blood types.
- (2) The surgeon is culpable for her ignorance only if she is culpable for the fact that she did not double-check the patient's chart.
- (3) We cannot establish with certainty that the surgeon is culpable for the latter.
- (4) Therefore: we cannot establish with certainty that the surgeon is culpable for the wrong blood transfusion.

We like to hold surgeons responsible in these kinds of situations. However, if Rosen is right, this is more difficult than it seems. If you wish to deny (4), and we assume again that (1) and (2) are plausible, then (3) must be called into question.

But Rosen is not so easily proved mistaken. Is the surgeon actually culpable for the fact that she did not double-check the patient's chart?

There are two situations. Either the surgeon is forgetful and forgets this more often or she is not forgetful. If she is not forgetful, then it was unreasonable to expect her to ask a colleague to help her remember that she had to double-check the chart, and then we should hold her responsible for forgetting this one time.

If the surgeon does often forget things, it would have been reasonable to expect her to ask the help of a colleague (or of some other kind of reminder). In that case, however, the question arises again as to whether she is culpable for not doing this or that she can perhaps be excused on the basis of ignorance.

In short, we are faced here with a regress in which we constantly have to check something else. We began with the question: Is the surgeon culpable for not double-checking the chart? And now: Is she culpable for negligence with respect to her forgetfulness?

On the one hand, if the surgeon was aware that she should do something about her forgetfulness but did not do anything, then she is culpable. But, Rosen argues, it is difficult to determine she did know this (if she already knew it beforehand). On the other hand, if the surgeon was not aware that she had to do something, then the question arises as to whether she was culpable for her ignorance (which takes us further into regress).

All in all, (3) seems to be plausible: we cannot establish with certainty that the surgeon is culpable for the fact that she did not double-check the patient's chart.

//

Compare this, however, to a different example:

- (1) The slaveholder is culpable for the exploitation of her slaves only if she is culpable for her ignorance of the view that slavery is morally impermissible.
- (2) The slaveholder is culpable for her ignorance only if she is culpable for the fact that she did not research the question better.
- (3) We cannot determine with certainty that the slaveholder is culpable for the latter.
- (4) Therefore: we cannot determine with certainty that the slaveholder is culpable for the exploitation of her slaves.

Moody-Adams would disagree with (4).<sup>29</sup> According to her, the ignorance of slaveholders is “affected” (in short, they benefit from maintaining their ignorance), and we can still hold them responsible.

Assignment 8.3:

If (4) is false, at least one of the premises must be false. Which of Rosen’s premises would Moody-Adams deny?

//

Assignment 8.4:

Evaluate your argument from assignment 2.4 regarding its soundness (so not just validity).

---

<sup>29</sup> Moody-Adams, M. M. 1994. Culture, Responsibility, and Affected Ignorance. *Ethics* 104.

## 9. Fallacies

Making a good argument is not the same as *persuading* others.

Persuading your audience is seeing to it that they accept your conclusion. But, in principle, it does not make any difference if you do that via a sound argument or via a fallacy

On the other hand, you can have a sound argument even though your audience does not accept your conclusion. (Though one should of course be careful here: if no one accepts your conclusion, including the experts who have thought about for a long time, then you should look at it again.)

Fallacies are arguments that could perhaps persuade but are not valid. Your task is to expose them.

Below is a series of common formal fallacies.

Ad verecundiam: appeal to authority

- (1) An irrelevant authority says that p is true
- (2) Therefore: p

Ad ignorantiam: appeal to ignorance

- (1) Nobody has proved not-p
- (2) Therefore: p

Shifting of the burden of proof

- (1) You have given no reason for not-p
- (2) Therefore: p

Ad baculum: appeal to the stick

- (1) Not accepting p has disadvantageous consequences for you
- (2) Therefore: p

Ad consequentiam: directed at consequences

- (1) Not-p has disadvantageous consequences
- (2) Therefore: p

The last two differ. Compare:

“If you don’t believe that this is a brilliant textbook, I will give your exam an unsatisfactory; therefore, this is a brilliant textbook.”

“If this is not a brilliant textbook, then the students will have difficulty with the rest of the study; therefore, this is a brilliant textbook.”

Ad populum: appeal to the people/common knowledge

- (1) Everyone knows that p is true
- (2) Therefore: p

Ad hominem: personal attacks

- (1) Not-p adherents are stupid (or lazy, ugly, etc.)
- (2) Therefore: p

Tu quoque: you too

- (1) You also accept p
- (2) Therefore: p

Please note: variations on this can sometimes produce good arguments:

- (1) You also accept p
- (2) But if p, then q
- (3) Therefore: q

Example:

“You also accept that exploitation is morally wrong; but if exploitation is morally wrong, then buying products that are made using exploitation is also morally wrong; therefore, buying those products is morally wrong.”

Ad misericordiam: appeal to sympathy

- (1) I have gone to great trouble to defend p
- (2) Therefore: p

There are other kinds of fallacies. For example, the invalid argument forms (affirmation of the consequent, negation of the antecedent) can also be misused to deceive an audience.

In general terms, you can assume that you are faced with a fallacy when it does not deal with the content of the conclusion and thus presents no information about the truth of that conclusion (but rather provides information about the hearer or something else that has no implications for truth or plausibility).

One could say that the tacit premise “if (1) then (2)” is implausible.

The following fallacy is, however, an exception to this characterisation:

Petitio principii: circular reasoning

- (1) p
- (2) Therefore: p

In this case, the premise gives information about the truth of the conclusion (and the tacit premise “if p, then p” is plausible). Nevertheless, this is seen as a fallacy because what has to be proven is already accepted (and this can mislead your audience).

Examples are not always immediately recognisable as such perhaps:

“Keeping slaves is against the law, therefore keeping slaves is punishable.”

This is a circular argument insofar as “being against the law” and “being punishable” are terms for the same property.<sup>30</sup>

Assignment 9.1:

Identify the fallacies in [this video](#) (in the first minute).

---

<sup>30</sup> Such arguments take the following form: (1) a is F; (2) F = G; (3), therefore, a is G. Given that this is not exactly the same as “p because p,” we need to be careful and not simply put all arguments with this form aside.

## 10. Rewriting

Most people cannot write—not even an email, not to mention a structured argument in which they defend a standpoint.

Writing is not simply seeing to it that you have said everything you want to say, stringing a bunch of sentences together.

Writing is also seeing to it that all sentences are relevant and that they are in the right place (namely, that they gradually move toward a standpoint).

Your task is learn to do that better.

Assignment 10.1:

Consider the following fragments. Use the tools you have become acquainted with so far and rewrite the texts.

If determinism is true, if, that is, all events obey immutable laws, then my will too is always determined, by my innate character and my motives. Hence my decisions are necessary, not free. But if so, then I am not responsible for my acts, for I would be accountable for them only if I could do something about the way my decisions went; but I can do nothing about it, since they proceed with necessity from my character and the motives. And I have made neither, and have no power over them: the motives come from without, and my character is the necessary product of the innate tendencies and the external influences which have been effective during my lifetime. Thus determinism and moral responsibility are incompatible. Moral responsibility presupposes freedom, that is, exemption from causality. (Schlick, *Problems of Ethics*, p. 146<sup>31</sup>)

First make a reconstruction of the argument:

- (1) Determinism is true. [ass]
- (2) Therefore: my decisions are completely determined by my character and motives.
- (3) I have not made either of them and have no power over them.
- (4) Therefore: I cannot exercise any influence on my decisions.
- (5) I am responsible for my acts only if I can exercise influence on my decisions.
- (6) Therefore: I am not responsible for my own acts.
- (7) Therefore: if determinism is true, I am not responsible for my acts.

You have to make choices in such a reconstruction process. E.g. I have ignored the term “freedom”, which complicates the argument unnecessarily. This term is explained in the last sentence (“exemption from causality”), and in the reconstruction I assume that this term is covered by “exercising influence.”

Tip:

---

<sup>31</sup> This is an argument that Schlick only mentions in order to show that it rests on confusion.

Use as little jargon (unusual terms or ordinary terms with unusual meanings) as possible. If necessary, define your terms (“freedom” in this case) *before* you present your argument, not at the end.

The rewritten text:

Thesis: if determinism is true, then I am not responsible for my acts. Suppose determinism is true. In that case, my decisions are completely determined by my character and motives. I have not made either of these, nor do I have any power over them. It follows that I can exercise no influence over my decisions. But such influence is necessary for responsibility. Therefore, if determinism is true, I am not responsible for my acts.

//

The difference between things that are temporarily assumed (assumptions), things that are considered to be true (premises), and things that are derived from other things (inferences) is key.

Explaining (1)-(7) above:

- (1) is an assumption that is only accepted to see what follows from it (i.e. (6)).
- (3) and (5) are premises.
- (6) is an inference from (4) and (5) via modus tollens.
- (7) is an inference from (1) and (6) via conditional proof.

That leaves (2) and (4). How does (2) follow from (1), and (4) from (2) and (3)?

None of the argument forms we saw in §4 seem to correspond to this.

Tip:

If you cannot move from A to B via one of the valid argument forms, try to see if an extra premise “if A, then B” is plausible and deduce B via modus ponens.<sup>32</sup>

This is possible in the case of (2) and (4):

- (2) can be derived from (1) via the extra premise “if determinism is true, then my decisions are completely determined by my character and motives” and modus ponens.
- (4) can be derived from (2) and (3) via the extra premise “if my decisions are completely determined by my character and motives and if I have not made either or have no power over them, then I cannot exercise any influence on my decisions” and modus ponens.

//

Antithesis: There is no absolutely necessary being existing anywhere, either in the world or outside the world as its cause.

Proof: Suppose that either the world itself is a necessary being or that there is such a being in it; then in the series of its alterations

---

<sup>32</sup> The argument form cannot itself be inserted as an extra premise. In the case of modus ponens, you cannot e.g. insert the following: ‘if p and ‘if p, then q,’ then q’. Why not? See: Carroll, L. 1895. What the Tortoise Said to Achilles. *Mind* 4.

either there would be a beginning that is unconditionally necessary, and hence without a cause, which conflicts with the dynamic law of the determination of all appearances in time; or else the series itself would be without any beginning, and, although contingent and conditioned in all its parts, it would nevertheless be absolutely necessary and unconditioned as a whole, which contradicts itself, because the existence of a multiplicity cannot be necessary if no single part of it possesses an existence necessary in itself... Thus neither in the world nor outside it (yet in causal connection with it) is there any absolutely necessary being. (Kant, *Critique of Pure Reason*, A453/B481, trans. P. Guyer<sup>33</sup>)

Make a reconstruction first:

- (1) Suppose something exists that is necessary. [ass]
- (2) Then it is either (a) the beginning of the series of changes in the world or (b) the series of changes itself.
- (3) Horn (a) yields a contradiction (namely, with the law that everything has a cause).
- (4) Horn (b) also yields a contradiction (because a whole cannot have the property in question if none of the parts have it).
- (5) Therefore: nothing necessary exists.

(1) is an assumption that is accepted only to derive a contradiction and ultimately to reject it. Thus, the general argument is a *reductio ad absurdum*.

But there is more. The sub-argument is a dilemma, namely between two horns that both lead to a contradiction.

The passage from Kant is structured very well. The most important task in rewriting is to simplify the language and drop terms like “conditioned,” “contingent,” and “absolute” that do not seem to add anything, at least not in this passage).

The rewritten text:

Does a necessary being exist? There are two possibilities. Either the beginning of the series of changes is necessary, or the series of changes themselves is necessary. The first option is implausible. For, if the beginning of the series is necessary, then it has no origin, and this contradicts the law that everything has an origin. The second option is also implausible. For, if the series itself is conditioned, whereas none of its parts is, then the whole would have a property that none of its parts have. In short, both horns lead to a contradiction, from which we can conclude that no necessary being exists.

//

Assignment 10.2:

---

<sup>33</sup> Kant defends not only the antithesis but also the thesis, namely, that there is an “absolutely necessary being.” Together the thesis and antithesis form an “antinomy”.



Choose a challenging passage with a high argumentative density of ~200 words from a famous philosopher (e.g. Kant) and rewrite this passage (i.e. based on your reconstruction of the argument).<sup>34</sup>

Choose, preferably, a text in which you think the philosopher makes a good point but does not state it as well as it could be stated.

You rewrite this text for your fellow students in this course.

Submit the original text, your reconstruction of the argument, and the rewritten text. Explain briefly what choices you have made during the process (in terms of the criteria of interpretation, charity, and simplicity).

---

<sup>34</sup> As in 2.4, you may consult *Just the Arguments*.

## 11. Counterexamples

X is a necessary condition for Y:  
Y *only if* X.

X is a sufficient condition for Y:  
Y *if* X.

The difference is crucial:

If X is sufficient for Y, and you hear that X is the case, then you also know that Y is the case (via modus ponens).

You do not know the latter if X is simply necessary for Y. After all, if X is the case, but another necessary condition for Y does not obtain, then Y does not obtain.

If X is necessary for Y, and you hear that X is not the case, then you also know that Y is not the case (via modus tollens).

You do not know the latter if X is only sufficient for Y. After all, there can still be other sufficient conditions for Y (and therefore Y).

In short, if it is assumed that certain things are necessary or sufficient, and you receive further information, then you also know the things that must follow from that. If they do not follow, then the things you started with are apparently neither necessary nor sufficient.

This is actually the *essence* of thought experiments.

//

The following are two well-known examples of thought experiments.

Frankfurt:<sup>35</sup>

Suppose someone—Black, let us say—wants Jones to perform a certain action.<sup>36</sup> 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 he 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. Whatever Jones's initial preferences and inclinations, then, Black will have his way.

What steps will Black take, if he believes he must take steps, in order to ensure that Jones decides and acts as he wishes? Anyone with a theory concerning what "could have done otherwise" means may answer this question for himself by describing whatever

---

<sup>35</sup> Frankfurt, H. G. 1969. Alternate Possibilities and Moral Responsibility. *Journal of Philosophy* 66.

<sup>36</sup> E.g. Black wants Jones to get rid of someone.

measures he would regard as sufficient to guarantee that, in the relevant sense, Jones cannot do otherwise. Let Black pronounce a terrible threat, and in this way both force Jones to perform the desired action and prevent him from performing a forbidden one. Let Black give Jones a potion, or put him under hypnosis, and in some such way as these generate in Jones an irresistible inner compulsion to perform the act Black wants performed and to avoid others. Or let Black manipulate the minute processes of Jones's brain and nervous system in some more direct way, so that causal forces running in and out of his synapses and along the poor man's nerves determine that he chooses to act and that he does act in the one way and not in any other. Given any conditions under which it will be maintained that Jones cannot do otherwise, in other words, let Black bring it about that those conditions prevail. The structure of the example is flexible enough, I think, to find a way around any charge of irrelevance by accommodating the doctrine on which the charge is based.

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, or to withhold the praise to which it would normally entitle him, 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. He would have acted the same even if it had not been a fact. Indeed, everything happened just as it would have happened without Black's presence in the situation and without his readiness to intrude into it. (1969: 835-6)

Gettier:<sup>37</sup>

Suppose that Smith and Jones have applied for a certain job. And suppose that Smith has strong evidence for the following conjunctive proposition:

- (d) Jones is the man who will get the job, and Jones has ten coins in his pocket.

Smith's evidence for (d) might be that the president of the company assured him that Jones would in the end be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago. Proposition (d) entails:

- (e) The man who will get the job has ten coins in his pocket.

Let us suppose that Smith sees the entailment from (d) to (e), and accepts (e) on the grounds of (d), for which he has strong evidence. In this case, Smith is clearly justified in believing that (e) is true.

But imagine, further, that unknown to Smith, he himself, not Jones, will get the job. And, also, unknown to Smith, he himself has ten coins in his pocket. Proposition (e) is then true, though

---

<sup>37</sup> Gettier, E. 1963. Is Justified True Belief Knowledge? *Analysis* 23.

proposition (d), from which Smith inferred (e), is false. In our example, then, all of the following are true: (i) (e) is true, (ii) Smith believes that (e) is true, and (iii) Smith is justified in believing that (e) is true. But it is equally clear that Smith does not know that (e) is true; for (e) is true in virtue of the number of coins in Smith's pocket, while Smith does not know how many coins are in Smith's pocket, and bases his belief in (e) on a count of the coins in Jones's pocket, whom he falsely believes to be the man who will get the job. (1963: 122)

These thought experiments constitute counterexamples.

Counterexamples can call a necessary or a sufficient condition into question.

**X is sufficient for Y:  
Y is always present if X is.**

**Counterexample:  
X without Y.**

After all, if we have a X without Y, then Y is not always present if X is and X is insufficient for Y.

Example:

Someone claims that 10 beers are sufficient for a hangover. Thus, if you drink 10 beers, you should have a hangover.

You produce a counterexample by finding a case in which someone drinks 10 beers but does not have a hangover.

**X is necessary for Y:  
Y is present only if X is.  
Or: X is present if Y is.**

**Counterexample:  
Y without X**

After all, if we have a Y without X, then X is not always present if Y is, and then X is not necessary for Y.

Example:

Someone claims that 5 beers is necessary for a hangover. Thus, if you have a hangover, you will have had 5 beers.

You produce a counterexample by finding a case in which someone has a hangover without having had 5 beers.

//

Assignment 11.1:

Explain whether the condition Gettier and Frankfurt call into question is a necessary or a sufficient one.

Gettier's case:

- there is JTB (“justified true belief”)
- no knowledge

This is a counterexample to:  
JTB is sufficient for knowledge.

Note: Gettier cases thus do not call into question any necessary condition for knowledge (check for yourself).

Argument:

- (1) Smith has JTB but has no knowledge.
- (2) Therefore: JTB is insufficient for knowledge.

Schematically:

- (1) X and  $\neg Y$
- (2) Therefore: X is insufficient for Y

Please note: (2) you can read this in two ways:

- If there is JTB, then there is no knowledge.
- It is not the case that: if there is JTB, then knowledge.

The first says that JTB and knowledge never go together. That is too strong. After all, Gettier cases only indicate that there are a few cases where this is so. Thus, only the second follows.

//

Frankfurt’s case:

- there is moral responsibility
- but no alternative possibilities of action

This is a counterexample to:  
Alternative possibilities for action are necessary for moral responsibility.

Note: Frankfurt cases thus do not call any sufficient condition for moral responsibility in question (again, check for yourself).

Argument:

- (1) Jones is morally responsible but has no alternative possibilities.
- (2) Therefore: alternative possibilities for action are not necessary for moral responsibility.

Schematically:

- (1)  $\neg X$  and Y
- (2) Therefore: X is not necessary for Y

(2) has two possible readings:

- If Jones is morally responsible, then Jones has no alternative possibilities to action.
- It is not the case that if Jones is morally responsible he then has alternative possibilities.

The first says that moral responsibility never goes together with alternative possibilities. Again, that is too strong. After all, Frankfurt's case constitutes one where that is so. Only the second follows.

//

The step to (2) seems to be deductively valid. But: how does (2) actually follow from (1) through propositional logic?

Instruction 1:

Rewrite the argument first in the language of propositional logic.

Key:

p: Jones is morally responsible.

q: Jones has alternative possibilities.

(1) is easy now:

$p \wedge \neg q$

(2) is more difficult:

$\neg(p \rightarrow q)$

These translations are wrong (check for yourself):

$p \rightarrow \neg q$

$\neg(p \leftarrow q)$

Instruction 2:

You now know that you want to conclude " $\neg(p \rightarrow q)$ ."

You can do this by accepting " $p \rightarrow q$ ," then attempt to deduce a contradiction and conclude " $\neg(p \rightarrow q)$ " through  $I_{\neg}$ .

- |    |                         |                          |
|----|-------------------------|--------------------------|
| 1. | $p \wedge \neg q$       | prem                     |
| 2. | $p \rightarrow q$       | ass                      |
| 3. | p                       | $E_{\wedge}$ (1)         |
| 4. | q                       | $E_{\rightarrow}$ (2, 3) |
| 5. | $\neg q$                | $E_{\wedge}$ (1)         |
| 6. | $q \wedge \neg q$       | $I_{\wedge}$ (4, 5)      |
| 7. | $\neg(p \rightarrow q)$ | $I_{\neg}$ (2, 6)        |

//

Note:

If X is necessary for Y, then Y is sufficient for X.

Example:

If alternative possibilities are necessary for moral responsibility, then moral responsibility is sufficient for alternative possibilities.

And conversely:

If X is sufficient for Y, then Y is necessary for X.

Example:

If JTB is sufficient for knowledge, then knowledge is necessary for JTB.

//

Assignment 11.2:

Choose a thought experiment (from a philosophical discussion that interests you), and discuss which necessary or sufficient condition is being questioned.<sup>38</sup>

Submit the original description of the experiment as well (or offer a good explanation of the experiment).

Some online experiments:

<https://www.philosophyexperiments.com>

//

Please note: not all thought experiments are presented straightforwardly as counterexamples and are more open-ended.

Consider Nozick's thought experiment:<sup>39</sup>

Suppose there were an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences? If you are worried about missing out on desirable experiences, we can suppose that business enterprises have researched thoroughly the lives of many others. You can pick and choose from their large library or smorgasbord of such experiences, selecting your life's experiences for, say, the next two years. After two years have passed, you will have ten minutes or ten hours out of the tank, to select the experiences of your next two years. Of course, while in the tank you won't know that you're there; you'll think it's all actually happening. Others can also plug in to have the experiences they want, so there's no need to stay unplugged to serve them. (Ignore problems such as who will service the machines if everyone plugs in.) Would you plug in? What else can matter to us, other than how our lives feel from the inside? Nor should you refrain because of the few moments of distress between the moment you've decided and the moment you're plugged. What's a few moments of distress compared to a lifetime of bliss (if that's what you choose), and why feel any distress at all if your decision is the best one? (1974: 42-3)

This thought experiment is often seen as a counterexample to hedonism.<sup>40</sup> After all, so goes the argument, if one can choose real life above an

---

<sup>38</sup> For a collection of thought experiments, cf. Tittle, P. 2004. *What If. Collected Thought Experiments in Philosophy*. Routledge.

<sup>39</sup> Nozick, R. 1974. The Experience Machine. In *Anarchy, State, and Utopia*, pp. 42-5. Basic Books.

<sup>40</sup> Against which view exactly: X is to be preferred above (or better than) Y *if or only if* X produces more pleasure than Y?

experience machine (where your pleasure is much greater than in true life), then pleasure cannot be the highest good.

Moreover, after this passage, Nozick presents various reasons why, according to him, one should indeed choose real life.

Nonetheless, the thought experiment can also be seen as an open-ended experiment and one in which the reader herself is invited to reflect on what is more important.



## 12. Definitions

Example of a definition:

“S is bored by X iff

- (i) S is in a certain mental state in which:
- (ii) S is weary,
- (iii) S is restless,
- (iv) S lacks interest in X to which she is subjected,
- (v) S’s state of weariness, restlessness, and lack of interest in X is unpleasant or undesirable to S, and
- (vi) S’s feeling of weariness and restlessness is causally connected to S’s lack of interest in X.”<sup>41</sup>

Please note: “iff” is not a spelling error; it is the abbreviation of “if and only if.”

Definitions present necessary and sufficient conditions. According to O’Brien, (i)-(vi) are necessary individually and sufficient collectively for boredom.

Another example:

“S knows that p iff

- (i) [that<sup>42</sup>] p is true,
- (ii) S believes that p, and
- (iii) S is justified in believing that p.” (Gettier 1963: 121)

Gettier does *not* subscribe to this definition. He presents it only to question it.

General formula of an definition:

X iff [Y, Z, ...]

whereby:

X is a philosophical concept (knowledge, responsibility, etc.) and Y, Z, etc. are individually necessary and collectively sufficient for X.

//

Definitions can be of different natures and have different functions.

The following distinctions are relevant:

- lexical vs. stipulative
- revisionist vs. descriptive
- intensional vs. extensional
- clarification vs. recognition

---

<sup>41</sup> O’Brien, W. 2014. Boredom. *Analysis* 74.

<sup>42</sup> Gettier makes a mistake in notation. Either add ‘that’ consistently (so that ‘p’ must always be filled in by a complete sentence) or leave ‘that’ out consistently (so that ‘p’ refers to a proposition). Note; if you use symbols, provide a translation key for how the symbols are to be read, and apply your translation key consistently.

A *lexical* definition is the definition of how a word is actually used or should be used and that can be found in a dictionary.

A *stipulative* definition is the definition of how the author wants to use a word, and that can deviate from how other people use it.

Within philosophy, many definitions are stipulative, and dictionaries are not very helpful. In everyday language use, concepts are often too vague to be of use in constructing interesting arguments (e.g. “freedom” or “causality”).

You could e.g. write:

“In this argument, I use the term causality in the counterfactual sense. Thus, by ‘X is a cause of Y’ I mean that Y had not occurred if X would not have occurred.”

As a philosopher, you can be more or less revisionist. A *revisionist* definition revises the current use of a concept, whereas a *descriptive* definition tries to stay closer to the current use.

//

In general, definitions indicate what the things that fall under the definition have in common. Thus, a definition of “moral permissibility” indicates what all morally permissible acts have in common.

Definitions can, however, focus on the core or commonality (the meaning), or rather the things that fall under it (the reference). An *intensional* definition sets the meaning of a word, while an *extensional* definition establishes the reference.

Example: “Dutch person”

Intensional definition:

A person with a Dutch history, culture, language, and ancestry.

This sets the meaning, though not immediately the domain of things that fall under it.

Extensional definition:

Anyone who is found within the borders of the Netherlands.

This definition is of course extensionally incorrect: not everything that falls under the definition also falls under the term “Dutch person” (e.g. a person who is on holidays in the Netherlands).

Listing everything that falls under the term also yields an extensional definition:

The collection of Jan, Marie, etc.

//

A definition can then have two functions: *clarification* or *recognition*.

If a definition is clarifying, it then helps us understand a term.

If a definition has a recognition function, it then helps to indicate whether something falls under a certain term.

The clarifying definition:

The above intensional definition does not immediately clarify matters. After all, the term “Dutch” is used in the definition itself.<sup>43</sup>

A further clarification can consist of a specific cultural features, such as the fact the people there love skating more than the average person.

The recognising definition:

People with a Dutch passport or who wear orange clothes. These are signs by which you can recognise a Dutch person.

There is a link between intensional definitions and the clarification function, as well as one between extensional definitions and the recognising function, but the labels are in principle independent.

//

A definition by Moody-Adams:

Affected ignorance—choosing not to know what one can and should know—is a complex phenomenon, but sometimes it simply involves refusing to consider whether some practice in which one participates might be wrong. Sometimes—perhaps much of the time—cultures are perpetuated by human beings who are uncritically committed to the internal perspective on the way of life they hope to preserve. (1994: 296)

Example from Moody-Adams:

The ancient Greeks were ignorant about the moral impermissibility of slavery. According to Moody-Adams, this was a question of “affected” (hypocritical) ignorance. They did not know better, but they did not question it either.

Another example:

Many consumers are ignorant of the circumstances in which the products they buy are made (clothing, food, electronics, etc.). Nevertheless, there is continually more information about this. It also seems to obtain here that consumers can do some research, but do not do so.

Moody-Adams’ definition (the first sentence of the quote) presented in terms of necessary and sufficient conditions:

S’s ignorance concerning p is “affected” iff

- (i) S chooses not to be better informed about p;
- (ii) S should be better informed; and
- (iii) S is also able to do it.

Examples of ‘p’:

- The slaves I hold are human beings.

---

<sup>43</sup> This yields a so-called “circular definition.” Circular definitions are not always problematic. This depends on what you want your definition to provide (e.g. clarification or recognition).

- The product I bought is made by people who are being exploited.
- This piece of food has been a living organism.

Type of definition:

Intensional rather than extensional: it establishes the meaning of the term rather than the domain to which it refers.

Assignment 12.1:

Discuss whether this is a good intensional or extensional definition, i.e. whether certain conditions are superfluous or missing.

Issues that you can ask include:

- S is ignorant, but should S not have a vague suspicion that p could be true in order to avoid such information?
- Is S's ignorance always embedded in a social context? (Think of the example of slavery.)
- Is there something general to say about S's motive? Does she benefit from her ignorance? Is the practice in which she participates advantageous for her? Does knowledge about p affect her self-image?

//

Some philosophy students assume that examples are superfluous, that texts can be written without even mentioning a single example. It is true that philosophers are inclined towards abstraction, and it is true that one can get lost in the details of an example. Nonetheless, without applications, general definitions and standpoints are meaningless.

Assignment 12.2:

Check if your conditions from the previous assignment cover all the examples by Moody-Adams below. You may also modify them if needed:

In practice, affected ignorance takes several forms; I discuss only four important varieties. The elaborate linguistic deceptions by which torturers are known to mask the reality of their activities illustrate a particularly malevolent variety of affected ignorance. Reports from around the world reveal a striking similarity in the way in which those engaged in torture describe their violent methods by means of deceptively benign phrases such as “the telephone” and the “parrots’ swing.” Such descriptions ultimately allow the torturer to deny the connection between his wrongdoing and the suffering of his victim.

To understand the second variety of affected ignorance, we can imagine the head of an investment banking firm who demands that her employees increase the firm’s profits but insists on knowing nothing about the means used to accomplish this. This executive’s wish to “know nothing” of the potential wrongdoing of her employees is surely—in some degree—culpable.

A third variety of affected ignorance is typically manifested in the readiness of some people to “ask no questions” about some state of affairs, in spite of evidence that an inquiry may be needed in order to stop or prevent wrongdoing. Thus a mother who repeatedly accepts expensive gifts from a teenage son with a modest income is

surely complicit in her son's wrongdoing—at least morally—if the gifts have been purchased with money from the sale of drugs.

Finally, perhaps the most common form of affected ignorance is the tendency to avoid acknowledging our human fallibility: as finite and fallible beings, even our most deeply held convictions may be wrong. But it is also common for human beings to avoid or deny this possibility. Mill relies on the prevalence of this variety of affected ignorance and on its common connection with the desire to suppress the convictions of others, in order to argue for freedom of thought and expression. Of course, Mill may be incorrect in his claim that all silencing of discussion is “an assumption of infallibility.” But he is clearly right to assume that there is a common human tendency to avoid, or even to refuse, acknowledgment of our own fallibility. Equally important, this tendency is often manifested in some of the most vicious wrongdoing possible. An extremely bigoted person who would violently silence protest of his bigotry is almost always guilty of this variety of affected ignorance. But it is at work even in less actively malevolent cases. A university administrator who refuses to investigate charges of wrongdoing because his colleague “couldn't possibly” be guilty of sexual harassment also manifests this kind of affected ignorance. (1994: 301-2)

Issues you can raise:

- Why should they be better informed?
- Can they be better informed?
- In what sense do they participate in something impermissible?
- What are their precise reasons for remaining ignorant?

//

Usefulness:

Define the terms of the position that you are defending. It is important to know what you are talking about.

And that your arguments are not ambiguous. A simple example:

- (1) Consequentialism condemns A if A makes a difference.
- (2) Nelson Mandela makes a difference.
- (3) Therefore: consequentialism condemns Nelson Mandela.

Such arguments seem to be valid, but only because the phrase “making a difference” is not defined. If you do define it, then the phrase acquires two different meanings in (1) and (2), and (3) no longer follows.<sup>44</sup>

Plus, definitions are sometimes interesting in themselves. People usually do not know what a certain concept (responsibility, freedom, identity, knowledge, etc.) means precisely and philosophers can offer clarification.

---

<sup>44</sup> E.g. an action A makes a difference with respect to an outcome O iff A is necessary for O (i.e. if S does not do A, then O won't occur).

### 13. Thought experiments I

In experiments, certain factors are isolated in order to investigate the relation between them. This also applies to thought experiments, even though this isolation takes place in the imagination and not in the world.

Examples:

In Gettier cases, we want to examine the relation between JTB and knowledge, and we isolate these factors in our imagination.

In Frankfurt's cases, we want to know what the relationship is between alternative possibilities and responsibility, and we isolate these factors in the imagination.

But the question is what this means concretely. What factors do we need to keep in an experiment, and what should we remove? How do we find suitable cases? (Or: How do you become famous like Frankfurt and Gettier?)

Thought experiments are not found by chance. There are general instructions that you can follow.<sup>45</sup>

The instructions depend on whether you want to undermine a necessary or sufficient condition. In both cases, there are four steps.

Here is the idea in brief (in terms of the example from §11):

Suppose someone states: 10 beers are necessary to have a hangover. What is the counterexample? A hangover without having 5 beers. But how do you keep the hangover? By adding a sufficient condition for the hangover and e.g. drinking a lot of vodka.

Suppose someone states: 10 beers is sufficient for a hangover. What does the counterexample look like? 10 beers without a hangover. But how do you remove the hangover? By removing a necessary condition for the hangover (e.g. lack of fluids) and drinking a lot of water.

//

The instructions for necessary conditions in steps:

Step 1:

Identify the standpoint of your opponent:  
X is necessary for Y.

Frankfurt's opponent:

Alternative possibilities for action are necessary for moral responsibility.

Step 2:

Where you want to end up: Y without X. Therefore:  
Retain Y and remove X.

---

<sup>45</sup> Wieland, J. W. & M. Endt 2017. Analysing Thought Experiments. *Teaching Philosophy* 40.

In Frankfurt's case:  
Retain moral responsibility and remove alternative possibilities.

The following step is crucial. The question is: how do you retain Y (moral responsibility)?

Step 3:  
Add another factor that you think is sufficient for Y.

In Frankfurt's thought experiments, the factor that Jones acts on the basis of his own will and reasons is added (or emphasised more).

This is a factor that seems to be sufficient for moral responsibility. And, thus, if you add that factor, you retain moral responsibility.

Step 4:  
Finally, add enough details to the case description so that the factors are interpreted correctly.

In Frankfurt's case:  
To ensure that the factor "no alternative possibilities for action" is interpreted correctly, you can add information about Doctor Black.

Other information (about Black's hobbies, Jones' friends, etc.) does not need to be added).<sup>46</sup>

//

The instructions for sufficient conditions:

Step 1:  
Identify the position of your opponent:  
X is sufficient for Y.

Gettier's opponent:  
JTB is sufficient for knowledge.

Step 2:  
Where you want to end up: X without Y. Therefore:  
Retain X and remove Y.

In Gettier's case:  
Retain JTB and remove knowledge.

The following step is crucial. The question: how do you remove Y (knowledge)?

Step 3:  
Remove another factor that you think is necessary for Y.

---

<sup>46</sup> The consequence is that thought experiments often come across as streamlined and unrealistic, but that does not have to be a problem (cf. also Fischer, J. M. 1995. *Stories*. *Midwest Studies in Philosophy* 20).

In Gettier's thought experiment, the factor that the grounds on whose basis Smith built his view are true is removed: they are false.<sup>47</sup> Smith believes that the man who gets the job has 10 coins in his pocket, but he believes this because he presupposes that Jones (not he) gets the job and has 10 coins in his trousers pocket, which is false.

True grounds seem to be a factor that is necessary for knowledge.<sup>48</sup> And thus, if you remove this factor, you remove knowledge.

Step 4:

Finally, add enough details to the case description so that the factors are interpreted correctly.

In Gettier's case:

To ensure that the factor JTB is interpreted correctly, add information about Jones (and in what way his view is true and justified).

//

X is necessary for Y iff Y is sufficient for X (see end §11).

Why *two* sets of instructions then?

If you wish to dispute a necessary condition for Y (knowledge, moral responsibility, or another concept that is analysed/defined), you need to retain Y in that case.

Frankfurt's cases are cases in which the agent *is* morally responsible.

If you want to dispute a sufficient condition for Y, then Y (the concept that is analysed/defined) does *not* need to be retained in the case.

Gettier cases are cases where the agent has *no* knowledge.

//

Assignment 13.1:

Choose two thought experiments from Thomson (1976, cited in §3) and check how they are constructed by means of the instructions just given above.

Below are two examples worked out in detail. So, choose further cases (there are at least 24 different cases in that paper, involving trolleys, surgeons, avalanches, atomic bombs, and health-pebbles).

The Passenger:

Frank is a passenger on a trolley whose driver has just shouted that the trolley's brakes have failed, and who then died of the shock. On the track ahead are five people; the banks are so steep that they will

---

<sup>47</sup> Clarke, M. 1963. Knowledge and Grounds. A Comment on Mr. Gettier's Paper. *Analysis* 24.

<sup>48</sup> It is of course a matter of dispute as to whether true grounds are indeed necessary; you can test this with further thought experiments.



not be able to get off the track in time. The track has a spur leading off to the right, and Frank can turn the trolley onto it. Unfortunately there is one person on the right hand track. Frank can turn the trolley, killing the one; or he can refrain from turning the trolley, letting the five die. (1976: 207)

Frank has a choice between not intervening (letting five people die) and intervening (killing one person). According to Thomson, Frank may intervene in this situation, which calls into question the theory that killing is always worse than letting die (a theory that is used against the permissibility of e.g. euthanasia).

Please note:

Thomson says that Frank *may* intervene, not that he *should*, or that *you* should do the same. For now, we will assume that Thomson is right about this.

Here the question is: how is a case like this constructed?

Step 1:

In terms of trolleys (and other cases Thomson discusses) we can formulate the theory of Thomson's opponent as follows:

T1 It is morally permissible to avert a threat *only if* this does not involve killing people.

Step 2:

Because we have to do with a necessary condition, we have to leave the factor "it is morally permissible to avert a threat" in the case and remove the factor "this does not involve killing people."

Step 3:

To retain "it is morally permissible to avert a threat" we must add another factor that seems to be sufficient for this. In this case, Thomson adds the factor "the new target involves fewer victims than the current target."

Step 4:

Add details about these factors, such as the trolley, the passenger, and the two tracks with people on them.

You could now think that Thomson adheres to the following alternative view:

T2 It is morally permissible to avert a threat *if* the new target involves fewer victims than the current target.

After all, in contrast to T1, T2 means that Frank can steer the trolley to the track with 1 victim instead of 5.

But Thomson does not defend T2 either, partly on the basis of the following thought experiment.

The Passer-by:

George is on a footbridge over the trolley tracks. He knows trolleys, and can see that the one approaching the bridge is out of control. On the track behind the bridge there are five people; the banks are so steep that they will not be able to get off the track in time. George knows that the only way to stop an out-of-control trolley is to drop a very heavy weight into its path. But the only available, sufficiently heavy weight is a fat man, also watching the trolley from the footbridge. George can shove the fat man onto the track in the path of the trolley, killing the fat man; or he can refrain from doing this, letting the five die. (1976: 207-8)

George is faced with the choice of not intervening (letting five people die) or intervening (killing one person). In this respect, the choice is the same as for Frank. Nonetheless, according to Thomson, George may not intervene in this situation (he is not allowed to throw the passer-by from the bridge). Yet T2 permits intervention if the new target involves fewer victims.<sup>49</sup>

How is this case constructed?

Step 1:

The theory to be criticised: T2.

Step 2:

Because we have to do with a sufficient condition, we have to leave the factor “the new target involves fewer victims than the current target” in the case and remove the factor “it is morally permissible to avert the threat.”

Step 3:

To remove the factor “it is morally permissible to avert a threat” from the case, we need to remove another factor that seems to be necessary for this.

In this case, Thomson removes the factor “you can avert a threat by acting against the threat itself, rather than against a person.” This factor differentiates between two ways of intervening. One can act directly against the threat (such as steering the trolley in a different direction) or act by using another person (such as throwing some passer-by in front of the train to stop it). And the latter is impermissible according to Thomson.

Step 4:

Finally, add details about these factors, such as the trolley, the passer-by who can stop the train, and the track with people on it.

//

Assignment 13.2:

Take your thought experiment from assignment 11.2 and see how it is structured in accordance with the above-mentioned instructions.

---

<sup>49</sup> Once again: we will set aside any questions we might have about Thomson’s intuitions.

## 14. Thought experiments II

T: the theory that X is necessary or sufficient for Y

P: verdict about a case

Argumentation scheme (modus tollens):

- (1) If T, then P.
- (2) P is absurd/false.
- (3) Therefore: T is false.

T: Alternative possibilities are necessary for moral responsibility.

Frankfurt:

- (1) If T, then Jones is not morally responsible.
- (2) Jones is morally responsible.
- (3) Therefore:  $\neg T$ .

T: Maximalisation of survivors is sufficient for moral permissibility.

Thomson:

- (1) If T, then George may sacrifice the passer-by.
- (2) George may not sacrifice the passer-by.
- (3) Therefore:  $\neg T$ .

What is the usefulness of such reconstructions?

You can see immediately what options you have if you want to defend T against these arguments: attack (1) and/or (2) (one of the two suffices).

Attack (1):

Define the theory in question (T) in such a way that a certain verdict about a case (P) no longer follows from T.

Example:

You could explain maximalisation in such a way that it concerns not only the maximalisation of survivors but also the consequences in society that such an intervention entails. Then, it might not simply follow that you may intervene if you are a consequentialist.

Attack (2):

Deny the judgment (or intuition) of the author regarding the thought experiment. This is also called "biting the bullet".

Perhaps you think that Frank may *not* divert the trolley from the five to the one. Or perhaps you think that George is allowed to throw the passer-by from the bridge.

That is possible. But only stating the opposite is not very interesting in philosophy. Why should others take your intuition seriously? (And why should *you* do that?)

It is more interesting to see if you can find an argument as to *why* Thomson is wrong.

What would such an argument look like? There are various options:

(a)

Construct a comparable case—i.e. with the same factors—that actually supports your intuition.

(b)

Conduct a poll among various groups (various cultures or socio-economic classes) and show that you are not the only one to have different intuitions than Thomson's.

This option cannot be carried out from the so-called “armchair” and is known as “experimental philosophy.”<sup>50</sup>

What if you know that 80% of people agree with Thomson (suppose)? You can ask yourself where that brings us. The question whether Thomson is right is, after all, not a matter of “the most votes.”

Indeed, experimental philosophy is useful in a different way. What polls can contribute is determining whether philosophers are not operating too much in a “no-man’s land”.

More research is of course necessary to see *why* people agree or disagree. Sometimes they can have good reasons, and those reasons can then be made explicit.

Their intuitions can also emerge from conceptual confusions (because they do not know the philosophical debates), or they can be the result of arbitrary and unreliable processes. This brings us to the third option:

(c)

Demonstrate that Thomson's intuitions are the result of arbitrary and unreliable processes.

Again, this cannot be done from your armchair, and you will have to consult the empirical sciences (e.g. psychology, neuroscience).

E.g. empirical studies have shown that reactions to the Passenger correlate with a heightened brain activity in the higher cognitive areas, whereas reactions to the Passer-by correlate with a heightened brain activity in the emotional areas of the brain.<sup>51</sup>

This would suggest that they concern different types of cases. The Passenger is an abstract dilemma, whereas the Passer-by is a concrete individual who must be thrown in front of the train.

Important: this is not to say that there is not a moral difference *as well* between the two cases (that acting in the one case is permitted and in the other not), but we need to do more work to show this.

(d)

---

<sup>50</sup> Knobe, J. & S. Nichols 2007. An Experimental Philosophy Manifesto. In *Experimental Philosophy*. OUP.

<sup>51</sup> Greene, J. et al. 2001. An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science* 293.

Show that there is an alternative theory that supports your deviating intuition (more on this in §15).

//

Assignment 14.1:

Most likely, you do not share all of Thomson's intuitions (or you know people who do not). Check if the disagreement can be supported by one of the above strategies.

Then check if the thought experiment in question can be improved to circumvent the objection.

Cases can be "underdescribed", in which case relevant information must be added. Or they may be "overdescribed" in which case irrelevant or disruptive information must be removed. Rewrite the experiment in such cases.

//

Assignment 14.2:

Describe how Nozick (1974: 42-5, cited in §11) continually adapts and modifies his experience machine, so identify the various versions.

Then read: De Brigard, F. 2010. If You Like It, Does It Matter If It's Real? *Philosophical Psychology* 23.

Which of the above-mentioned strategies does De Brigard use against Nozick, and do you find this convincing?

## 15. Reflective equilibrium

Suppose someone presents a thought experiment, and asks for your response by posing the question to you: “May Frank intervene?”

How would you actually answer? May you take some time to reflect? More specifically, are you allowed to resort to theories to give an answer?

You might think you are, but that yields a problem.

At first, intuitions seem to be there to undermine certain theories and support other theories. Thus, the intuition that Frank can intervene in the Passenger argues against T1 and for T2 (see §13).

But if this is the role intuitions play, one should not use theories to respond to thought experiments. You then become caught in a circle.

To state the problem more precisely:

- Are theories plausible because they correspond to intuitions in specific cases?
- Or are intuitions in specific cases plausible because they correspond to the theories?

Example:

- Is T2 plausible because it agrees with intuitions about the Passenger?
- Or are intuitions about the Passenger plausible because they agree with T2?

This problem is known as “the problem of the criterion,” and it plays a central role in ancient scepticism.

The sceptics saw two answers to the problem:<sup>52</sup>

Put all intuitions and theories on hold, and stop with one’s inquiry (go do something else).

Put all intuitions and theories on hold, and proceed (forever) with one’s inquiry.

These sceptical answers are not that popular today. Philosophers usually want to have recourse to theories. Those who do only ask questions without ever finding any answer and philosophers who stop asking questions completely are viewed with suspicion.

There are three non-sceptical alternatives:<sup>53</sup>

Particularism:

Intuitions are primary. Theories must conform to intuitions, not the other way around.

Methodism:

---

<sup>52</sup> The second option is known as “Pyrrhonism.”

<sup>53</sup> Chisholm, R. M. 1973. *The Problem of the Criterion*. Marquette UP.

Theories are primary, and intuitions have to conform to theories, not the other way around.

Reflective equilibrium:

Both intuitions about specific cases and general theories are important, and must be weighed against each other.

Most philosophers—implicitly or explicitly—follow the last option (background [here](#)).

They will admit that theories can be revised in the light of intuitions regarding specific cases (as we saw in Thomson 1976).

But also that intuitions (or judgments about specific cases) can be revised in the light of a theory.<sup>54</sup>

The theory must of course have certain advantages to fulfill this role. The theory must e.g. be able to account for many other cases. It also helps if the theory is more elegant (simpler) than alternative theories, and if it can give a uniform explanation for its verdicts about specific cases.

For the answers to thought experiments, this means that you can reflect as long as you want and that you can consult theories while doing this.

Finally, you can be more or less revisionist. That is, you can try to link up as closely as possible to everyday intuitions, or you can challenge these intuitions and question them.

E.g. Derek Parfit counts as a revisionist:<sup>55</sup>

Descriptive philosophy gives reasons for what we instinctively assume, and explains and justifies the unchanging central core in our beliefs about ourselves, and the world we inhabit. I have great respect for descriptive philosophy. But, by temperament, I am a revisionist. ... I try to challenge what we assume. Philosophers should not only interpret our beliefs; when they are false, they should *change* them.

//

Consider the following alternative for T1 and T2:

- T3 It is morally permissible to avert a threat if
- (i) the new target involves fewer victims than the current target;
  - and
  - (ii) you can intervene by taking action directly on the threat rather than a person.

This theory is based on the following suggestion by Thomson:

---

<sup>54</sup> This is option (d) from §14.

<sup>55</sup> From the introduction to *Reasons and Persons* (OUP, 1984).

what matters in these cases in which a threat is to be distributed is whether the agent distributes it by doing something to it, or whether he distributes it by doing something to a person. (1976: 216)

Assignment 15.1:

Check if T3 covers both the Passenger and the Passer-by.

Then draw up a table in which you indicate what the theories say about these cases (whether you may or may not intervene in them), and note if that is correct according to Thomson.

	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>Thomson's intuition</i>
Passenger				
Passer-by				
...				

Assignment 15.2:

Discuss the extent to which T3 is plausible. Specifically, consider further cases she discusses (in subsequent work such as: Thomson, J. J. 1985. The Trolley Problem. *Yale Law Journal* 94) and note in the table when T3 gives the intuitively correct analysis and when it does not.

Finally: try to formulate a theory that could cover all cases.

//

Philosophers do not often use tables or other graphic material (diagrams, schemas, etc.). Nevertheless, these devices can help to clarify the issue at hand at a glance (e.g. whether a theory has counterexamples). Train yourself to check if your text can be supported in this way.



## 16. Research

Types of questions (part 1): conceptual, empirical, normative.

Conceptual questions:

- What is knowledge?
- What is strategic ignorance?
- What is poverty?

Empirical questions:

- How much poverty exists in the world?
- Why do people maintain strategic ignorance?
- Is climate change actually happening?

Normative questions:

- Is knowledge valuable?
- Are we morally obligated to fight poverty?
- Is killing always worse than letting die?

Philosophers usually look at conceptual and normative questions and bracket empirical questions (which can be researched in other sciences).

After all, those questions are precisely one can deal with via their methods (such as thought experiments).

Assignment 16.1:

Formulate a conceptual question, an empirical question, and a normative question about the covid crisis.

Possible source:

Bramble, B. 2020. *Pandemic Ethics. 8 Big Questions of COVID-19.*  
Bartleby.

Please note:

Be careful with empirical information. The interpretation of empirical studies on e.g. climate change is not always easy, and requires expertise.

Should you nevertheless want to refer to empirical studies (e.g. to then pose a normative question about climate change), you can best introduce them as assumptions instead of premises (see §4 for the difference).

//

Types of questions (part 2): systematic, historical, applied.

Systematic question:

- Is there a good solution for the counterexamples to Kant's formula of universal law?

Historical question:

- What exactly were Hegel's objections to Kant's formula of universal law?

Applied question:

- Is slavery to be condemned on the basis of Kant's formula of universal law?

//

Usually, interesting questions are too large to tackle in one go (unfortunately), and you need to properly delimit your research question.

There is a fixed formula for this:

"Based on argument Y, author this or that holds that X is the case; is that a good argument?"

Example:

Based on his analogy with the child in the pond, Singer argues that we should do more to fight poverty; is that a good argument?

Assignment 16.2:

Formulate three well-delineated research questions that philosophers can ask about [see syllabus].

//

Research into such questions is done along established lines:

- What is Singer's standpoint precisely?
- What is his argument? What are the premises?
- Does the standpoint follow from the premises?
- Are the premises plausible, or are there counterarguments or counterexamples that can be conceived?

Based on such research, you can take a standpoint *yourself*.

Types of standpoints:

- Interpretation: you argue that Singer's standpoint can be read or elaborated in different ways or that a straightforward reading does not make sense.
- Validity: you argue that Singer's argument is invalid, or that a certain tacit premise is implausible.
- Soundness: you argue that a certain premise Singer holds is implausible, or, in contrast, you provide additional support for it.
- Application: you argue that Singer's position or argument has an unexpected or interesting application.

Assignment 16.3:

Formulate three possible standpoints that you could take on the basis of research into the questions in 16.2.

## 17. Writing

Writing in philosophy is not just writing some text but arguing towards a certain conclusion.

What you will learn here: **to write based on an argument you have thought out beforehand and have composed *before* you start writing.**

The reader of an argument is not interested in your thought process that led to the argument. Rather, the reader wants to read your final argument.

Suppose your thought process goes as follows:  
A therefore B, because C.

Then your argument is apparently plural: you have *two* arguments for B: A and C. Thus, do not write “A therefore B, because C,” but rather “C because A and B.”<sup>56</sup>

Advantage: if your argument is already complete, you won’t be faced with “writer’s block.” For you already know what you want to write about.

That does not, however, mean that your text has to be boring, that you simply put down the various steps of your argument one after the other. Even if your argument is set, you still have enough artistic freedom to write a fine piece.

//

Consider the following cases:<sup>57</sup>

### *Tourist*

Will is a tourist, and when he is about to leave for his holiday destination a stranger offers him \$2000 to deliver a suitcase abroad. The stranger looks friendly, and although Will suspects that something might be wrong with the contents of the suitcase, he keeps himself willfully ignorant by not asking any questions. He tells himself that it might well contain gifts for the stranger’s family and accepts the offer. Unknown to him, the suitcase contains weapons. Noah is very similar to Will, and receives and accepts the same offer, though in contrast to Will he does ask about the contents of the suitcase, and knows full well he is transporting weapons.

### *Conductor*

Will lives during World War II and helps the Nazis as the conductor of a transport train. He doesn’t know about the Holocaust, because he keeps himself willfully ignorant by not asking any questions. He suspects that something might be wrong about his work, but he doesn’t know that he transports people to concentration camps. Noah is very similar to Will, and does the same kind of work for the Nazis. In contrast to Will, however, Noah doesn’t fool himself. He

---

<sup>56</sup> Alternatively, your argument is simple, and A and C *together* form an argument for B. This difference matters: in the first case, a proponent of C must deny both A and B; in the second, denying one of the two suffices.

<sup>57</sup> Wieland, J. W. 2019. Willful Ignorance and Bad Motives. *Erkenntnis* 84.

asks about the passengers, and knows full well he's transporting people to concentration camps.

### *Flyer*

Will is a frequent flyer. Every weekend, he flies to another city to spend his free time abroad. Of course, he has heard people make claims about climate change, but he doesn't really know whether they are true since he keeps himself willfully ignorant by not reading any newspapers and not looking into joining a carbon offset program. He suspects that something might be wrong about his lifestyle, but he doesn't know that he contributes to climate change. Noah is very similar to Will, and leads a very similar party life, though unlike Will he knows full well about his ecological footprint.

So there are two individuals: Will and Noah. In all cases, they do precisely the same, but there is this difference: Noah knows what he is doing, whereas that does not hold for Will. Will does have some suspicion that there could be something wrong, and he knows that he could be better informed but decides not to do that.<sup>58</sup>

The question now is: who can we hold more responsible?

Four positions:

- (a) Will is more culpable than Noah.
- (b) Noah is more culpable than Will.
- (c) They are equally culpable.
- (d) It depends, namely on ...

Assignment 17.1:

Choose one of the answers, and defend it on the basis of your own reasoning in the form of an argument (250-300 words).

Example 1:<sup>59</sup>

Position (b) seems the most acceptable to me. Noah can be reproached for more than Will. In connection with this, I think that Noah and Will can be held accountable in two different ways. That is to say, if I say that Noah is more culpable for something than Will, I qualify what it concerns as "more serious" than the other thing (for which Will is culpable). Thus, the scale whereby Noah ends on the more evil side is not the same as the one used for Will.

Will keeps himself ignorant despite the fact that he, with a bit more information, might have made a different choice. The reasons for this could be more diverse but seem to have the common denominator that Will thinks that he needs to adapt and that his situation could be worse. He can be reproached for this.

Noah is not to be reproached for this. He does not allow himself to be led by the possible inconvenience that awaits him when he knows what is actually going on. But Noah can be held accountable for something else, namely, that he has made a (reasonably) informed choice, which he also knows will cause harm.

---

<sup>58</sup> His ignorance is strategic or "affected" (a concept discussed in §12).

<sup>59</sup> Both examples are written by students in class.

This is a different sort of reproach but one that I view as more serious than that for which we can hold Will accountable.

Example 2:

In all three of the given situations, Will chooses not to know what is going on. He thus maintains his ignorance. That for which he is partly to be held accountable is simply the fact that he deliberately discards any responsibility in order to avoid further complications. He sees to it that he himself will remain free of blame. His culpability thus consists in that he sets up a limit of knowledge and does not want to recognise other consequences out of self-interest. Will is thus very egocentric.

Nevertheless, this claim is context-dependent because he is egocentric and lazy in scenario 3, but that does not have to be the case in scenarios 1 and 2. His motives for being ignorant are based in fear. Perhaps he does not want to be involved in crimes against humanity in scenario 2 because his resistance makes little difference. Perhaps he has family that he has to maintain with his job as a train driver. There are extra factors that also play a role in his choice for ignorance.

Noah is convinced, however, of or indifferent to his acts. If he is convinced, then he can be held accountable for much. He acts wrongly and he also sometimes also defends his action. If he is indifferent, the same factors can play a role as with Will. Then they can only be reproached for being passive, but with the good reasons there can even be understanding for that. The good thing for both is that if they had been in the right circumstances, they can be perhaps be equally morally competent. Further context is needed to give a correct evaluation.

Assignment 17.2:

Give feedback on these examples. Identify both good points and points that could be improved.

Think about the following possible comments:

- What do you mean? Explain.
- Difficult to follow. Rewrite.
- Too complicated. Simplify.
- Do you need this term? Avoid jargon.
- Too abstract. Example?
- Why should the reader believe this? Give better arguments.
- Too ambitious. Limit your point.
- How does this follow?
- What thus follows? Finish your argument.
- What are you doing here? Link to the preceding?
- Does not seem to be relevant.
- Unnecessary repetition. You've already said this.
- Where do you get this from? Give a precise reference.
- Why is this here? Move it somewhere else.

//

Assignment 17.3:

- Make a plan for an essay, where you defend a particular standpoint in the philosophical debate on [see syllabus].
- Make this plan according to the “four-sentence structure” (see below).
- Use 1 page for this plan, where you briefly clarify and defend each of the four sentences, especially sentence 2.
- Formulate your standpoint on the basis of the instructions in §16.
- Literature: use at least one source listed in the syllabus. You may consult additional sources, but in principle you can complete this assignment by using a limited number of discussion partners and by thinking for yourself.

**The four-sentence structure:**<sup>60</sup>

1. **[debate] Author X says that \_\_\_\_\_**
2. **[standpoint+argument] I say that \_\_\_\_\_ because \_\_\_\_\_**
3. **[discussion] You could object that \_\_\_\_\_**
4. **[reply] I answer that \_\_\_\_\_**

Example:

1. Rosen says that slaveholders are excused on the basis on ignorance.
2. I say that that excuse does not hold water because their ignorance is rooted in self-interest.
3. You could object that they did not know that they acted out of self-interest.
4. I answer that that is not relevant.

Important:

A good standpoint is interesting, but not too ambitious. Thus: *not* “free will does not exist.” Why not? The debate on free will has a long history, and it is impossible to discuss all proposed definitions of free will and to refute all defences of it.

It is better to make one limited point and to work it out in detail (explain it, defend it, illustrate it) than to make many major points that you cannot defend. Science—and philosophy—advances in small steps.

Please note: making a good plan usually costs more work than writing the essay itself. Take the time for this.

Assignment 17.4:<sup>61</sup>

- Write out your plan from 17.3 in 800-1000 words (devote about the same amount of words to each of the four sentences).
- You should situate your standpoint in the debate and start the discussion with others, though do not summarise others extensively. Use most words to write your own argument.
- Write from the “I” perspective: *you* have something to tell the reader about the debate. That is not to say that you need to act as if you are going to solve the debate. Keep your contribution limited, and

---

<sup>60</sup> Earl, D. 2015. The Four-Sentence Paper: A Template for Considering Objections and Replies. *Teaching Philosophy* 38.

<sup>61</sup> Further useful instructions: “[Guidelines on writing a philosophy paper](#)” or “[A brief guide to writing the philosophy paper](#)”.

make a modest point (and make the limitations of your argument explicit).

- Audience: write as if you are writing for fellow students who have not taken this course (and so do not know the debate).
- Use as little jargon as possible. Sometimes you need to introduce philosophical terms. If so, explain them to your audience. Never write a sentence that you would never speak yourself, and never write something that you yourself do not understand!
- See to it that *every* sentence and *every* word in your essay is relevant. Sentences should have a clear message: giving either new information or a new step in your argument supporting your standpoint.
- Do not use footnotes but references (with page numbers) and a literature list (according to the instructions in §19).
- Add separately (this does not count for the amount of words): a reconstruction of your argument in steps, and an explanation of its validity.

//

### Assessment

Minimum requirements (to receive a grade):

- carrying out the assignment
- meeting the deadline
- within the word count
- correct grammar/spelling
- correct referencing and bibliography
- no plagiarism

Rubric:<sup>62</sup>

<b>Criterion</b>	<b>Good [1 point]</b>	<b>Satisfactory [0.5 point]</b>	<b>Unsatisfactory [0 point]</b>
(1) Standpoint	The standpoint is unambiguously stated.	The formulated standpoint is clear, but not all ambiguities have been clarified.	No clear standpoint has been formulated, or it must be distilled from the text by the reader.
(2) Argument (content)	The argument for the standpoint is well worked out and convincing.	The author presents a reasonably worked out argument, but is not entirely convincing.	The argument remains superficial (i.e. the reader has to finish the argument herself).
(3) Structure (form)	The argument is thought out well, and has a logically valid form: dilemma, reductio, etc.	The argument is thought out well, though does not have a logically valid form.	The argument is not thought out well.

<sup>62</sup> The latest version of this rubric is by Linda Holland.

(4) Focus (content)	The essay is properly demarcated, and the limitations of the argument are made explicit.	The essay is properly demarcated, but the limitations of the argument are not made explicit.	The essay is too ambitious and does not support the standpoint.
(5) Relevance (form)	The whole essay contributes to the argument (no needless repetition or digression).	The majority of the essay contributes to the argument (some needless repetition or digression).	The minority of the essay contributes to the argument (too much needless repetition or digression).
(6) Illustration	The argument is well illustrated by means of a suitable example of which the importance is made explicit.	An example is given, but it is not completely appropriate or its importance remains implicit.	No example is given, and therefore the argument remains abstract.
(7) Terminology	The philosophical terms are explained well, and the argument is written for the right audience.	Not all philosophical terms are explained well, but the argument can be followed by the audience in question.	Not all philosophical terms are explained well, and the argument cannot be followed by the audience in question.
(8) Literature	The author presents the literature adequately, and really engages it well.	The author presents the literature adequately, but does not fully engage it.	The author does not present the literature adequately, or does not engage it at all.
(9) Originality	The argument displays the author's own thinking, adds something to the literature and is thought provoking.	The argument displays the author's own thinking, and adds something to the literature.	The author reproduces the literature in a passive way, without any critical reflection or addition.

Grade: 1 + max. 1 point per criterion

//

#### Assignment 17.5:

Give the members of your group feedback on their concept essay using the rubric. Please indicate at least one good point and one point that needs improvement in the concept essay. Make sure that your feedback is clear for the authors.



## 18. Presentation

Presenting is not the same as writing. You do not have to present everything that you have written. Provide only the most interesting points.

Some tips:

- Start with the central problem. Why is your topic interesting? What is stimulating about it? Explain the problem to your audience.
- Keep your research plans or results simple. Your audience would like to understand what you are saying.
- But you are the expert; you—not they—have done the work. Make this clear by, among other things, indicating that there are other relevant details or issues that you cannot discuss in your presentation. At the end, indicate problems or open questions that still occupy you; perhaps others can help you.
- Make use of a good handout (or possibly, a powerpoint) with a reconstruction of your argument.
- As a general rule: reading presentations is a no go (unless you can read well).

What obtains for reconstructions of argument also applies to handouts and powerpoints: keep it simple and clear.

E.g. do not make a handout or slide of a quote like the following (seen earlier), but present only your reconstruction in steps.

“The problem, in effect, is this: consequentialism condemns my act only when my act makes a difference. But in the kind of cases we are imagining, my act makes no difference, and so cannot be condemned by consequentialism—even though it remains true that when enough such acts are performed the results are bad. Thus consequentialism fails to condemn my act.”

Assignment 18.1:

Consider [this talk](#) by Peter Singer (or some other TED talk by a philosopher) and identify some differences from academic texts.

//

Assignment 18.2:

Prepare a 5-minute presentation on your essay plan. Take the time limit seriously.

Target group: a group of fellow students.

Assignment 18.3:

Present your argument to your group. Give each other peer feedback, indicating both good points and points that need improvement, and report further questions or suggestions.

The goal of this assignment is not only to sharpen your own argument but also to develop your skill to *give* feedback. Good feedback is not easy and needs practice.

//

## “The ethics of argumentation”

Making arguments can sometimes come across as an aggressive enterprise. If you have taken a certain standpoint and others disagree with you, then they attempt to undermine your argument, and you are expected to defend your view. That can be daunting and leave a negative impression.

It is true that there are no “safe” standpoints within philosophy. If you take a position, you are vulnerable to criticism.

It is good first of all to remember that it can be more important to produce a good argument (that surprises, fascinates, and is difficult to refute) than to adopt a position. Your standpoint is only as good as your best argument for it. And if your argument does not appear to hold water, it should then not be a problem in principle to modify your position or to abandon it completely.

Nevertheless, arguing (like any other undertaking) can be done in a harsher or less harsh way, or be directed more or less at collaboration. Thus, keep some “norms” in mind when you give and receive feedback.

### Giving feedback:

- Direct your criticism at the argument, not the person.
- Apply the principle of charity: attempt to assess as strong a version as possible of the argument.
- Be careful and keep in mind that you can be wrong. Perhaps your criticism is based on a misunderstanding.
- Indicate it when there is lack of clarity, and possibly indicate how this can be resolved.

### Getting feedback:

- Do not take the criticism personally: it is your argument that is at the centre.
- See the criticism as a challenge and an opportunity to sharpen your argument (or else to modify your standpoint).
- If your argument generates a lack of clarity, then that is, as a rule, not the fault of those who give you feedback.
- It is fine if you do not have any immediate answer. Some counterarguments need to be allowed to sink in.

Finally, arguing is a discipline with quite a lot of freedom. There are of course clear rules on which arguments are valid and which are not, but there is a lot of room to work within your own values.<sup>63</sup>

---

<sup>63</sup> See also the blog [“Is polite philosophical discussion possible?”](#) by Nomy Arpaly on this issue.

## 19. Literature

If you defend a standpoint on a certain issue, you should situate it in the literature. Why?

First, it is pointless to propose something that is uninformed by what others have said about it.

Second, you can actually use existing work to find a position in the first place, or to develop and refine your own standpoint.

Third, if you want others in the debate to take your standpoint seriously, you have to show how your contribution is related to what has already been said.

But how do you find relevant texts?

1. A search with key words
2. Literature reviews on the topic
3. References in and to texts you already know

1.

Many texts can be found via Google Scholar:

<https://scholar.google.com>

Usually they are behind a “paywall,” but you can gain free access via the university.

If you nevertheless do not seem to have access via the university, click on “All versions” under a result and see if you do have access via another link.

You can also start a Google search. You sometimes find the paper on the author’s website, or you can request the paper from the author via email.

Which search terms?<sup>64</sup>

- topic: “strategic ignorance”
- theory: “formula of universal law”
- case: “experience machine”
- problem: “false positives”
- central phrase: “do I make a difference”
- central author: Kant

Language: search in English. After all, you are looking for an answer to your research question, and you will not find it if you ignore international sources.

Search also via synonyms. Authors sometimes use other terms (e.g. “willful ignorance” or “affected ignorance” instead of “strategic ignorance”).

If you place search terms between quotation marks, the phrase will appear as a whole in the text.

---

<sup>64</sup> You can also set up an automatic notification so that you receive an email if a new publication appears, i.e. that falls under the search term.

You will often have very many hits, and you will have to focus your search by introducing other search terms.

Example: “formula of universal law” “false positives”

In principle, texts can appear from *all* disciplines in your research. To ensure that you only get philosophical texts, you will have to add enough philosophical search terms or select philosophical journals (e.g. *Ethics*, *Mind*, *Philosophy & Public Affairs*, *Philosophical Studies*, etc.).

Some journals are better than others (an example of an influential ranking).

You can then screen the articles from these journals (based on the abstracts and key words), and check if they are indeed relevant to your own goals.

Finally, you can focus through filters like publication date. Sometimes you are only interested in the latest publications on a certain topic, and you can finetune the period (e.g. “since 2020”).

Please note: philosophical publications usually have less quotations than publications in many other disciplines. One reason for that is philosophers usually only cite works *that they have actually read* (and then the literature lists are automatically shorter).

2.

A literature review provides the state of affairs regarding research into a certain question.

Such reviews do not take a position, but describe developments in some area as a more neutral spectator.

This neutrality does not mean that all literature reviews are the same: you still have to make choices about which works you find important to discuss and how you see the coherence between the different sources.

*Philosophy Compass* is a journal that only publishes literature reviews.

Literature reviews can also be found in online encyclopedias:

Stanford Encyclopedia of Philosophy  
<https://plato.stanford.edu>

The Internet Encyclopedia of Philosophy  
<https://www.iep.utm.edu>

3.

Finally, take texts that you already know (and that you find interesting and good) and check:

- What sources does the text refer?
- What other sources refer to this text?

Below a search result in Google Scholar is the text “Cited by [number of texts]” and “Related Articles” that you can click on.

//

Assignment 19.1:

Make a literature list of 5 philosophical publications on [see syllabus].

Important: do not trace the whole debate, or list random publications, but limit your search as much as possible, and *focus on one particular issue*.

Give a short explanation of the selection of your sources. Why did you choose precisely these works? And: describe how they are connected. How do the various works respond to each other (explicitly or implicitly)?

//

Below are instructions for structuring a literature list (based on the so-called "APA system"). These instructions provide all information that is necessary to locate a text. In principle, you can also use other instructions (as long as you are consistent).

Article:

Kagan, S. 2011. Do I Make a Difference? *Philosophy & Public Affairs* 39: 105-41.

[last name], [initials] [year of publication]. [title of article]. [*journal*] [volume number]: [page numbers].

Book:

Parfit, D. 1984. *Reasons and Persons*. OUP.

[last name], [initials] [year of publication]. [*title of book*]. [publisher].

Translated work:

Kant, I. 1785. *Groundwork of the Metaphysics of Morals*. Trans. M. Gregor & J. Timmermann 1998. CUP.

[last name], [initials] [year of original publication]. [*translated title*]. Trans. [translators] [year of translation]. [publisher].

Note: don't write things like "Kant, I. 1998".

Chapter:

Parfit, D. 1984. Five Mistakes in Moral Mathematics. In *Reasons and Persons*, pp. 76-86. OUP.

[last name], [initials] [year of publication]. [title chapter]. In [editor, if present], [*title of book*], pp. [page numbers]. [publisher].

Online encyclopedia:

Rickless, S. 2015. Plato's *Parmenides*. In *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/plato-parmenides/>

[last name], [initials] [year of publication]. [title]. In [*encyclopedia*], [url]

//

Once you have found relevant sources, you can work them into your own text. You have different options:

- refer
- quote
- paraphrase

## Reference

If you wish to refer to a text, then place the last name of the author and the year of publication in parentheses, as well as the page numbers (if relevant).

## Example:

Consequentialism faces different problems, including the collective action problem (see Kagan 2011).

Provide the complete information on the source at the end of your paper in the literature list.

## Quoting

If you want to use sentences directly from a source in their entirety, you can quote these sentences by placing them between quotation marks and ending with a reference.

## Example:

Kagan poses the collective action problem as follows: “consequentialism condemns my act only when my act makes a difference. But in the kind of cases we are imagining, my act makes no difference, and so cannot be condemned by consequentialism.” (2011: 108)

If you omit the quotation marks, then it is as if it is your own reasoning and you are committing plagiarism: a no go.

## Paraphrase

If you paraphrase, you rewrite the argument of someone else in your own words. You do not need any quotation marks for this, but you do need to make a proper reference.

## Example:

According to consequentialism, an act is morally wrong only if there was an alternative act with a better result. When buying chicken, there is no alternative with a better result, and thus buying chicken should be morally permitted according to consequentialism. Intuitively, however, such a purchase is not permitted, and this thus poses a problem to consequentialism (see Kagan 2011: 108).

As a rule, paraphrasing is better than quoting. Formulations can always be improved, and you can implement improvements in your paraphrase. You can sometimes omit or summarise relevant information, and sometimes you need to add relevant information (see §10).

Quoting is thus a form of laziness: you are too lazy to make such improvements.<sup>65</sup>

---

<sup>65</sup> But: you can sometimes quote brilliant philosophers—if you then engage in a rigorous analysis of the quote.

## 20. Exam

General instructions for an exam, i.e. with open questions:

Tip 1:

Answer questions step by step so that the examiner can see that you understand the question.

Thus, do not simply put down as much information as possible with the hope that the right answer is somewhere to be found. Such answers are not given any points.

Tip 2:

Illustrate your answers with examples and give a short explanation of the philosophical terms you are using (such as “modus tollens”) so that the examiner can see what you want to say with them.

The same obtains for reconstructions of arguments, translations, arrow diagrams, etc. They do not speak for themselves; explain all choices you make.

*What does this mean, concretely?*

Example question:

Why is “affirmation of the consequent” deductively invalid?

*Answer the question in steps. You can receive points for each step.*

*Step 1*

Affirmation of the consequent: if  $p$ , then  $q$ ;  $q$ , therefore  $p$ .

*Step 2*

Example: if it's raining, the streets are wet; the streets are wet; therefore, it's raining.

*Step 3*

That an argument is deductively invalid means that the conclusion does not necessarily follow from the premises. In other words, a situation is conceivable in which the premises are true, but the conclusion is not.

*Step 4*

Suppose I have made the streets wet myself and the sun is shining. This is a situation in which the streets are wet and where the streets are wet if it rains, but it is not raining.

*Step 5*

Finally, conclude your answer:

Hence, a situation is conceivable in which the premises are true, but the conclusion is not, and therefore the affirmation of the consequent is deductively invalid.