

An Opinionated Area Review:

Intuitions and Values: Re-assessing the classical arguments against quantitative hedonism



DAVID LANIUS

Abstract

Few philosophers today embrace quantitative hedonism, which states that a person's well-being depends only on the amount of her experienced happiness and suffering. Despite recent attempts to rehabilitate it, most philosophers still consider it untenable. The most influential arguments levelled against it by Mill, Moore, Nozick and Kagan purport to demonstrate that well-being must depend on more than only the amount of experienced happiness and suffering. I argue in this paper that quantitative hedonism can rebut these arguments by pointing out a shared systematic flaw in their argumentative structure. In particular, I argue that they are based on thought experiments that invoke either structurally unreliable intuitions or intuitions that are not in tension to the tenets of quantitative hedonism. While this does not rehabilitate the theory by itself, it shows that the classical arguments against quantitative hedonism provide less evidence against it than commonly thought and certainly do not conclusively prove it wrong.

Keywords: *hedonism, thought experiment, experience machine, intuition pump, well-being*

Introduction

Rarely any philosopher nowadays embraces quantitative hedonism, which states that a person's well-being depends only on the amount of her experienced pleasure (or happiness) and displeasure (or suffering). This is astonishing, since historically it used to be an influential and powerful theory, both in antiquity and in the nineteenth century during the period of British empiricism.¹

Kriterion – Journal of Philosophy, AO, AO(AO): 1–32.
<http://www.kriterion-journal-of-philosophy.org>

Roger Crisp [12, pp. 619f.] offers three reasons to explain the decline of (quantitative) hedonism in the twentieth century.² First, many regarded it as a “philosophy of swine” including – very prominently – John Stuart Mill, whose attempt to adjust it by distinguishing between lower and higher pleasures has, however, been widely considered to be either incoherent or to abandon hedonism altogether. Second, G. E. Moore formulated several arguments against hedonism including the *Heap of Filth* thought experiment, which were at the time considered to seriously undermine hedonism. Third, Robert Nozick then dealt, as Crisp [12, p. 620] puts it, a “near-fatal blow” to hedonism with his famous *Experience Machine* thought experiment.

Although historically somewhat less influential, there is another argument in the same spirit that should be added to the list. Shelly Kagan came up with a thought experiment that is widely taken to be the best example of modern criticisms against hedonism. Kagan’s argument of the *Deceived Businessman* appears even more powerful than the earlier ones by Mill, Moore and Nozick.

Despite these arguments and the fact that hedonism has nearly vanished as a position in the debate, it doesn’t seem anymore so obvious that it is as “wildly implausible” as Nozick [45, pp. 42ff.] or Sobel [52, p. 244] have claimed.³ Recent attempts to rehabilitate hedonism show that as a theory of what a good life consists in (prudential or value hedonism) and as a basis for a theory of what is right to do (hedonistic utilitarianism) it is anything but untenable.⁴

The arguments leveled against hedonism purport to demonstrate that the amount of overall happiness cannot be the only factor when evaluating a person’s well-being or moral actions. They aim to show that we care about things besides our experiential states, and they do so by invoking intuitions – that we intuitively prefer a state, world or action for the sole reason of its beauty, autonomy, truth or, in the case of Mill, that it generates more noble pleasures. From this they draw the conclusion that the amount of pleasure is not the only thing that intrinsically contributes to well-being or the morality of actions.

In this paper I will argue that the arguments by Mill, Moore, Nozick and Kagan share the same systematic flaw. There are two inherent problems they have in common. First, they are based on thought experiments that for the most part invoke structurally unreliable intuitions. Second, with regard to the intuitions that can be considered reliable, what the thought experiments show is not in contradiction with or even in tension to hedonism.

I will briefly discuss the four thought experiments and reconstruct the corresponding arguments (section 1), before presenting a general argument scheme that explicates their common structure (section 2). I will then argue why the thought experiments are unreliable and ultimately unsuccessful in showing what they aim to show (section 3).

1 *The Thought Experiments*

The classical arguments made against hedonism involve thought experiments. They invoke intuitions that aim to show that we care about more than merely the amount of pleasure – and thus that quantitative hedonism is false. There are, of course, important differences between them. As I will argue, however, they have essentially the same structure. In this section I will reconstruct the four main arguments with special focus on the thought experiments they rely on – to make this common structure transparent. Naturally, such a reconstruction will not be uncontroversial. I think, however, that even though the authors themselves may have had different intentions with regard to their arguments, they typically are presented (in the literature, but also in many textbooks on ethics) in a way that makes most sense if understood in the way reconstructed here.

The goal will thus be a moderate one – I aim to show that hedonism can escape the conclusions of the classical arguments as they are typically presented in the debate and what premises have to be accepted to do so. Consequently, there may be other, more sophisticated, arguments (or versions of these arguments) that refute hedonism after all, and the premises that hedonism is committed to may still be unacceptable to many of us.

Note also that the arguments have been interpreted as being directed against prudential hedonism, value hedonism, and hedonistic utilitarianism respectively.⁵ My reconstruction of them is intended to show that this difference in aim makes no difference to their evaluation; all four arguments fail for the same reason. Keeping this in mind, please bear with me in the following reconstruction and assessment of the four classical arguments against hedonism!

1.1 *The Philosophy of Swine*

Jeremy Bentham [4, p. 206] famously said that, quantity of pleasure being equal, pushpin is as good as the arts and science of music and

poetry. Thomas Carlyle saw this position with “inveterate dislike” and considered it “a doctrine worthy only of swine”.⁶ Mill’s restatement of this criticism can reasonably be seen as a thought experiment in its own right.

Let us imagine the life of a fully satisfied pig wallowing in the mud.⁷ The pig experiences more pleasure (both in terms of duration and intensity) than any human being and certainly more than you and me. We can conceive of it as the most satisfied pig ever – as ever-happy.⁸ Would you trade places with it? Mill claimed:

Few human creatures would consent to be changed into any of the lower animals, for a promise of the fullest allowance of a beast’s pleasures; no intelligent human being would consent to be a fool [...]. It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. [42, sec. II 6]

Almost no one would trade places with the pig. Most of us have strong intuitions to reject the offer to become an ever-happy pig. The thought experiment seems to show that there are some pleasures that are so exquisite – like listening to the opera or reading Greek philosophers – that no amount of lesser pleasures could compensate them.⁹ Trading one’s life for the life of an ever-happy pig would still make one worse off, even though we, as human beings, may experience less intense and fewer pleasures. The reason lies, so it seems, in our ability to experience pleasures of a higher quality.

The argument of the *Philosophy of Swine* (S) can be restated in the following way:

- (P_S^i) Almost no one would trade their life for that of the ever-happy pig.
- (P_S^e) If almost no one would trade their life for that of the ever-happy pig, people value some pleasures intrinsically more than others.
- (P_S^r) If people value some pleasures intrinsically more than others, there is more to well-being than the amount and intensity of pleasure.
- (P_S^h) If hedonism is correct, well-being depends only on the amount and intensity of pleasure.

(C_S) Hedonism is false.

Mill's thought experiment has been used to show that hedonism cannot account for the intuition that we don't want to live the life of the ever-happy pig. We consider some pleasures more valuable than others because of some intrinsic quality – not merely due to them being more intense or longer-lasting.

The argument is deductively valid and contains *prima facie* plausible premises. Premise (P_S^i) is the premise relying on the *intuition* that is invoked with the thought experiment; namely that we don't want to trade our life for the one of the pig. Premise (P_S^e) is the *essential* premise used in the argument, which connects the intuition with a more general claim about what we value. Premise (P_S^r) is a general *reliability* claim that we don't systematically err with regard to values. Premise (P_S^h) is simply the premise characterizing *hedonism*. And, finally, the conclusion (C_S) states that hedonism is false.

1.2 *The Heap of Filth*

Historically, Moore was crucial in bringing hedonism down. His most devastating argument against it used the *Heap of Filth*, which Moore took to thoroughly refute the theory that happiness is the only intrinsic value. Hedonism is false because there is more we value than mere experiential states. He states the thought experiment in the following words:

Let us imagine one world exceedingly beautiful. Imagine it as beautiful as you can; put into it whatever on this earth you most admire—mountains, rivers, the sea; trees, and sunsets, stars and moon. Imagine these all combined in the most exquisite proportions, so that no one thing jars against another, but each contributes to the beauty of the whole. And then imagine the ugliest world you can possibly conceive. Imagine it simply one heap of filth, containing everything that is most disgusting to us, for whatever reason, and the whole, as far as may be, without one redeeming feature. [43, sec. 50]

There are two worlds – one beautiful and one ugly. And, crucially, there is nobody who will ever experience either of them. This given, Moore [43, sec. 50] asks us which of the two worlds has more value, whether “it is better that the beautiful world should exist than the one which is ugly”.

Moore’s thought experiment is different from Mill’s one in the sense that it does not involve the lives of sentient beings. We have to determine the value not of one’s life, but of a world independent of anyone experiencing it. It has been criticized extensively that it is not clear what it could mean for a world to be more or less valuable if there is nobody to experience it.¹⁰

We can, however, look at the action of creating one world or the other and ask the question Moore asked: “Would it not be well, in any case, to do what we could to produce [the beautiful world] rather than the [ugly one]?” Since we are here not only concerned with a person’s well-being, but also with the question of what ought to be done, Moore’s thought experiment can – in principle and in this formulation – pose a serious problem to hedonism. The argument of the *Heap of Filth* (H) can then be precisified as follows:

(P_H^i) Almost everybody would rather produce the beautiful world.

(P_H^e) If almost everybody would rather produce the beautiful world, there is more that people intrinsically value than mere experiential states.

(P_H^r) If there is more that people intrinsically value than mere experiential states, there are other intrinsic values than pleasure.

(P_H^h) If hedonism is correct, the only intrinsic value is pleasure.

(C_H) Hedonism is false.

Moore used his thought experiment to show that quantitative hedonism cannot account for the intuition that we value beauty, even if there is nobody to appreciate it. The argument, as reconstructed above, is deductively valid and contains prima facie plausible premises.

Premise (P_H^i) encapsulates the intuition that we like a world full of beauty more than one which is a “heap of filth”. Premise (P_H^e) establishes the connection between this intuition and the claim that we intrinsically value beauty, and premise (P_H^r) states that if we think it more valuable to create a beautiful world rather than an ugly one, there must be another intrinsic value – namely, beauty, which is independent from any experiential state. Thus, combined with premise (P_H^h), hedonism is false.

1.3 *The Experience Machine*

Virtually every single book on ethics that discusses hedonism rejects it using the *Experience Machine*. Undergraduate philosophy students are usually introduced to this thought experiment for the primary reason of being persuaded that hedonism is untenable.¹¹

Consider how Nozick originally stated it:

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. [45, pp. 42-43]¹²

We are further told not to worry about our family or possible breakdowns of the machine. We are assured that everyone we care about will also enter the machine (or otherwise be happy) and that the machine works perfectly. We can now choose to plug into the experience machine. Would you do it?

Most people will pass on a life of great pleasure in the experience machine for what they currently have. They will prefer their *real* and less pleasurable life to the artificial one in the experience machine. This shows – so Nozick – that more than just our experiential states matter because living in reality matters to us, too. Sobel [52, p. 244] considered the intuitions invoked by the experience machine to strike “at the heart of hedonism”.

The argument of the *Experience Machine* (E) can be formulated as follows:

- (P_E^i) Almost no one would enter the experience machine.
- (P_E^e) If almost no one would enter the experience machine, there is more that people intrinsically value than pleasure.
- (P_E^r) If there is more that people intrinsically value than pleasure, there are other intrinsic values than pleasure.
- (P_E^h) If hedonism is correct, the only intrinsic value is pleasure.

(C_E) Hedonism is false.

According to Nozick, his thought experiment shows that hedonism cannot account for the intuition that we do want our life to track the truth, even if it makes us less happy. His argument, as formulated here, is deductively valid and contains prima facie plausible premises. Premise (P_E^i) is the premise relying on the intuition that is invoked with the thought experiment; namely that we don't want to enter the experience machine. Premise (P_E^c) is the premise connecting this intuition with a more general claim about what we value. Most people don't want to enter the experience machine because they think that a life in it is not real and that experiences based on reality are more valuable than those created by a machine. In short, we don't want to enter a machine that makes us happy but disconnected from reality. Finally, premise (P_E^v) is the assumption that we can generally trust what people value to actually be valuable. That is, there must be some intrinsic value to connectedness to reality, authenticity or truth. But this entails, combined with premise (P_E^h), that hedonism cannot be correct.¹³

1.4 *The Deceived Businessman*

Hedonism claims that only the amount of pleasure is ultimately relevant when evaluating someone's well-being. So far this claim has been critiqued by pointing out that not all pleasures are equally relevant (in the *Philosophy of Swine* argument) or that there is more than the amount of pleasure – such as beauty or truth – that has intrinsic value (in the *Heap of Filth* and *Experience Machine* arguments).

Kagan's *Deceived Businessman* aims to show that there are pleasures of a certain kind – namely “false” pleasures – that are worth less than “true” pleasures.¹⁴ Kagan describes the thought experiment in the following words:

Imagine a man who dies contented, thinking he has achieved everything he wanted in life: his wife and family love him, he is a respected member of the community, and he has founded a successful business. Or so he thinks. In reality, however, he has been completely deceived: his wife cheated on him, his daughter and son were only nice to him so that they would be able to borrow the car, [etc.]. [29, p. 311]

Has the deceived businessman had a good life? Hedonism seems to be committed to say yes. However, almost everybody thinks that the deceived businessman has had a bad life. If we had the choice to live the

same life without being deceived, we would choose to be the businessman whose experienced pleasures were grounded in reality. The deceived businessman's pleasures are in some sense "false" or "fake" and are thus less valuable. The argument of the *Deceived Businessman* (D) goes as follows:

- (P_D^i) Almost no one considers the life of the deceived businessman as good as one of the businessman who was not deceived.
- (P_D^e) If almost no one considers the life of the deceived businessman to be as good as the one of the businessman who was not deceived, there is more that people intrinsically value than mere experiential states.
- (P_D^r) If there is more that people intrinsically value than mere experiential states, experienced pleasure cannot be all that matters to well-being.
- (P_D^h) If hedonism is correct, experienced pleasure is all that matters to well-being.

(C_D) Hedonism is false.

Kagan argues that hedonism cannot account for the intuition that the deceived businessman has had a bad – or at least worse – life.¹⁵ Premise (P_D^i) encapsulates this intuition. Premise (P_D^e) states that this intuition entails that we value other things besides experienced pleasure – such as tracking the truth, being connected to reality or simply not being deceived. If we intrinsically value truth, however, there is – according to premise (P_D^r) – in fact some intrinsic value besides experienced pleasure. Consequently, hedonism must be wrong because the lives of the businessman who was deceived and of the one who was not deceived are valuable to different degrees, even though the pleasure experienced is quantitatively equal in both. Kagan's argument is deductively valid and contains prima facie plausible premises.

2 *The Arguments' Common Structure*

The four arguments share the same argumentative structure. Starting with the assumption that hedonism implies that pleasure is the only intrinsic value, they invoke a thought experiment which shows that most

people think that there is more to well-being or the morality of actions than solely the amount of experienced pleasure. The argument scheme is a *reductio ad absurdum* to show that hedonism is false and it is a typical instance of the *Method of Cases*.¹⁶

It can be stated in this way:

(P^i) Faced with the thought experiment, most people react in way R .

(P^e) If most people react in way R , then they hold normative conviction *non-H*.

(C_1) Most people hold normative conviction *non-H*.

(P^r) It is not the case that most people err with regard to normative convictions.

(C_2) It is the case that *non-H*.

(P^h) If hedonism is correct, then it is the case that H .

(C) Hedonism is false.

We start with people’s reaction to the thought experiment (P^i). Most of us intuit that it would be a bad idea to become an ever-happy pig or to enter the experience machine. Why do we think so? Why do we consider the deceived businessman’s life to be worse? The fact that many people share these intuitions – that they react in the same way – must be explained.

The second step is an inference to the best explanation. It is argued (or often just claimed) that the best explanation for these shared intuitions is that some pleasures are more noble than others, some pleasures are “false” due to the way they are brought about or simply that people intrinsically value beauty or truth (P^e). By this inference we conclude that people hold such normative convictions (C_1).

Only in a third step, another conclusion is drawn, namely that there really are the corresponding intrinsic values – intrinsic values incompatible with hedonism (C_2). This conclusion depends on the general reliability claim that people don’t systematically err with respect to normative convictions (P^r).

Finally, since hedonism entails that there cannot be such intrinsic values (P^h), it must be false (C). Or so it seems.¹⁷ In the next section,

I will show what hedonism is committed to in light of the thought experiments and why the arguments by Mill, Moore, Nozick and Kagan, as they are presented here, do – at the end of the day – not pose a threat to it.

3 *The Thought Experiments' Implications*

The thought experiments function both as *alethic refuters* – in Roy Sorensen's sense – and as *intuitions pumps* – in Daniel Dennett's sense. As alethic refuters, they are intended to falsify a statement or theory by constituting a counterexample to it.¹⁸ As intuition pumps, they invoke reactions with strong intuitive components.¹⁹ While thought experiments are not problematic by and in themselves, our intuitive judgments concerning their ramifications cannot simply be relied upon without justification.²⁰ Our intuitions “crucially and unavoidably” influence our overall judgments, as Bostrom and Ord [5] claim, and psychological biases tend to distort our intuitive judgments when faced with thought experiments.

Such distortion is particularly likely if features are stipulated that are so unrealistic that we have not experienced anything like them – as it is the case in the thought experiments.²¹ Qua thought experiments, they require us to imagine something to be the case. We are asked to imagine a possible world in which everything is true that is entailed by the description of the thought experiment. But not any such possible world will count: the possible world must be closest to our actual world. Only then is it allowed to draw inferences from our intuitive judgments.

There is a methodological choice on a continuum between two extreme approaches to evaluating such thought experiments – when imagining possible worlds in order to pump moral intuitions. One extreme on the continuum is what I call *trust approach*: we rely on the description of the thought experiment and take everything stipulated at face value. The other extreme on the continuum is what I call *realism approach*: We rely on our previous knowledge about the world when entertaining the experiment and contextualize it according to what we already know and take to be plausible.

Of course, we usually do something in-between – and unconsciously so.²² We always rely on the description of the thought experiment and try to take into account the stipulations as adequately as possible (otherwise it wouldn't make sense to entertain the thought experiment in the first place). At the same time, we always fill in the details of the thought

experiment with our world knowledge and rely on what we consider to be plausible (after all, the whole point of a thought experiment is to invoke intuitions based on what we imagine to be the case).

In the following subsections, I will argue that this methodological choice makes it very hard for us to reliably infer anything from our intuitive reactions to the thought experiments.²³ There are three options for the hedonist to rebut the arguments. First, she can doubt premise (P^i) – maybe people don't find it as revolting to become an ever-happy pig or prefer as strongly the beautiful world as it is claimed. As a matter of fact, there have been empirical studies which cast doubt on the claim that intuitions are as clear and widespread as many in the debate have assumed.

I don't want to discuss the validity of such empirical studies in this paper.²⁴ I think that the hedonist should have an explanation for anyone who has reaction R in face of the thought experiment, even if at the end there are not as many as originally assumed. So let us take premise (P^i) for granted.

The second option consists in attacking premise (P^e). The hedonist can argue that there is a better explanation for the fact that most people react in way R than that they hold normative conviction *non-H*. It is not self-evident that reaction R includes or implies any judgment about values.

Finally, the third option is to attack premise (P^r), arguing that in the special circumstances of the thought experiments, people in fact err with regard to normative convictions. Even if people explicitly make value judgments when confronted with the thought experiments, they may be mistaken.

3.1 Between Trust and Realism

Keeping in mind that it is not opaque to us where our intuitions come from when entertaining thought experiments in which we can imagine different scenarios on a continuum between *trust* and *realism*,²⁵ we may revisit the four arguments and ask ourselves

- (1) whether our intuitive reaction when faced with the thought experiments really is conclusive evidence for the claim that we have *non-hedonistic normative convictions* – i.e. evidence for the inference to the best explanation in premise (P^e) – and
- (2) whether these convictions also reliably track *actual moral values*: Can we rely on our normative convictions when formed in the

context of the thought experiments – i.e. are we warranted in assuming the reliability premise (P^r)?

The Philosophy of Swine Revisited

The *Philosophy of Swine* argument has been heavily criticized very early on by, for instance, Moore [43], Green [22] and Bradley [6]. Hauskeller [25, p. 428] has more recently argued that Mill failed to defend utilitarianism against the (at that time) potentially fatal charge that it is a philosophy “worthy only of swine” because he tacitly presupposed non-hedonistic standards.

So, let’s look at the intuitions invoked by the thought experiment. Apparently, we don’t want to trade our life for that of the pig. But why? We may feel disgust when imagining the life of a pig – be it as happy as possible. Also, such a happy life seems extremely boring. But, evidently, our feelings of disgust or our sense of boredom when considering the life of the ever-happy pig must not influence our judgment on whether to trade our life for that of the pig. The pig will – by stipulation – not experience disgust or boredom.

While it is certainly a fact that some activities (such as listening to the opera) can produce more and more lasting pleasure for some people. It is also true that other people get more pleasure from watching football games than from listening to the opera. The pig will – by stipulation – be as happy as the happiest person listening to any opera can be. We, as humans, may not be able to really imagine to get as happy by wallowing in the mud as by listening to the opera because we have already experienced the “higher” pleasures of listening to the opera. They are “higher” only in the sense that having experienced them precludes us from experiencing some other – “lower” – pleasures. It doesn’t mean that the fact that some people get more and more lasting pleasure from listening to the opera is a reason to rate the pleasure experienced *objectively* higher.

Another reason which might prevent us from trading with the pig is that we would lose control about our life. Who does guarantee that we’d be and remain happy as a pig? From our experience we know that pigs are usually not very happy and, crucially, lack authority over their lives. Most pigs’ lives in the real world seem miserable. If something goes wrong, as a pig I have few means to ensure my happiness. The experimenter could change her mind or be hindered to ensure my everlasting happiness. Then, I would live the life of a pig in the normal

world – which seems horrible. If we are to trust the experimental design, however, this – by stipulation – cannot happen.

A more serious reason to not trade with the pig concerns our personal identity. Will we remain the same person when becoming the ever-happy pig? This is hard to imagine. But if we cannot imagine it, there is further reason not to trade with the pig, since we don't want to cease to exist. We wouldn't be the ones experiencing all the happiness that *this* pig is going to experience. But, then, it seems better to remain a slightly less happy human being – and to remain oneself.

As the persons we are, we intuit – for all these reasons – that we wouldn't be happy doing what the pig does and can do. Mill's thought experiment is strongly affected by *status quo bias* and our anthropocentric perspective – we assess the pig's life from our own (human) standards.²⁶ We cannot imagine the pig to be *truly* happy. That is, we aren't seriously entertaining the idea of the ever-happy pig and thus don't trust the experimental design. Then, however, premise () is not warranted because there are better explanations for most people's initial reactions than their purported conviction that the ever-happy pig's life is worse than a less happy human's one because pigs cannot experience as noble pleasures as humans can.

But even if we trust the experimental design, there is still strong reason not to become the ever-happy pig. For instance, as pigs, we cannot support and help other people. We would endanger our means to make other sentient beings happy. How many other persons could we – as happy pigs – make happy? This alone is a striking reason to reject becoming the ever-happy pig.

Let us now with best intentions and efforts try to ignore such additional intricacies and further stipulate that our becoming ever-happy pigs wouldn't affect anybody else. We may even assume that everyone will be happier than before. Then, the situation becomes less clear. Either we still have the inclination to not become pigs and stick to our intuition, or we – faithfully imagining the thought experiment – deflate it.

In the first case, I argue, there is some evidence that we are mistaken. Irrelevant factors such as the sense of boredom and disgust or the feeling of responsibility for other people tend to (subconsciously) sneak into the possible world that we imagine for the thought experiment.²⁷ If we do, however, these are better explanations for our intuitive reaction *R* than our having non-hedonistic normative convictions.

However, even if people consciously come to the belief that some

pleasures are intrinsically more valuable than others, there is reason to doubt that they established the belief in a sufficiently reliable way. Consider this alternative thought experiment which we can call *Alien Pleasure*:

Alien Pleasure

Imagine that we get in contact with a superior alien race with much better means to get the most noble pleasures. Should we trade our life for a life of such an alien, knowing that it will be less pleasurable in intensity and duration, but with the guarantee that we will experience more noble pleasures than we have ever experienced before?

If we completely trust the experimental design and choose to remain human in the *Philosophy of Swine*, as Millians we must choose to trade our lives for those of the aliens, since they are able to experience the nobler pleasures. If we are not willing to do it, we are only left with distrusting the experimental design or choosing to become the ever-happy pig in the *Philosophy of Swine*. Either way, the anti-hedonistic intuitions are deflated. If we choose to become the ever-happy pig, premise (P^e) is not warranted. If we distrust the experimental design, the normative convictions we gathered from entertaining the thought experiment are at best doubtful. If we don't faithfully imagine a possible world in which the pig is *truly* happy, premise (P^r) is not warranted. Why should we trust our intuitions when formed under biased conditions?

It is anything but clear where our intuitions come from – even when we choose to become aliens. We don't just know the normative convictions *non-H*. We bring information with us that might or might not belong to the thought experiment, such as a sense to miss something great – e.g. the noble pleasure of listening to alien opera – something that would outweigh additional (expected) happiness we as ordinary humans would experience.

Hedonism thus has the resources to counter the *Philosophy of Swine* argument by deflating the intuitions invoked by it – both on the level of normative convictions (P^e) and actual values (P^r). Our intuitive reaction R and our holding normative convictions *non-H* can best be explained with recourse to our cognitive limitations and the questionable reliability of our intuitions when imagining something as unrealistic as becoming the ever-happy pig.

The Heap of Filth Revisited

The *Heap of Filth* faces some of the same problems as the *Philosophy of Swine*, even though it was originally directed just against value hedonism. There are other important differences, too. The main objection that has been made against the *Heap of Filth* is that Moore inadvertently contaminated his worlds with consciousness.

It is close to impossible to discount *oneself* – to remove oneself from the equation – when entertaining thought experiments. And, as the persons we are, it naturally is better to have a beautiful world than to have an ugly one – because it brings more pleasure to experience beauty than to experience ugliness. Surely, we are supposed to imagine the worlds entirely empty of sentient observers. But how can we do such a thing?

Can we really exclude the possibility that some sentient being might somehow stumble on the world at some time? For this being it would be better to find it beautiful. But even if we could exclude this possibility, simply producing one of the two worlds gives us pleasure to different degrees. It arguably is more pleasurable to create a beautiful world. And, finally, if either world simply existed, it arguably is still more pleasurable to *imagine* the beautiful world.

We cannot think about the world without thinking about it from some perspective. But if we tacitly and inadvertently introduce some observer, premise (P^e) is not warranted. The best explanation for reaction R is not that people believe beauty to be intrinsically valuable. The best explanation is that people imagine a world that in some way contains or creates more happiness.

If we trust the experimental design, we have to exclude any conscious being – we have to exclude ourselves from the thought experiment. But does it then make sense to even ask the question anymore? How should anything have value – if not to *someone*?

It is nowadays rather uncontroversial that there is something structurally amiss in Moore's *Heap of Filth*.²⁸ If people, after becoming aware of these problems, still have reaction R when faced with the thought experiment, it cannot be inferred from this that they believe beauty to be intrinsically valuable. Even if people explicitly claim to believe beauty to be intrinsically valuable, the unrealistic nature of a comparison between such uninhabited worlds provides strong reason to doubt premise (P^r).

Thus, the same argumentative flaw appears in the *Heap of Filth*: We cannot trust our intuitions when doing something as outlandish as comparing Moore's worlds, since we (subconsciously and systematically)

introduce irrelevant factors into the equation: Either people don't actually hold normative convictions *non-h* or they are likely to err with respect to them.

The Experience Machine Revisited

Let us now consider the *Experience Machine* argument in more detail. There is a huge and ongoing debate on the setup of the thought experiment and the intuitions invoked by it. At least three basic problems have been identified. First, people resist imagining the very possibility of such a machine and worry, for instance, that it might not work as well as stipulated. Second, people find it hard to give up the responsibility for their loved ones. And, third, people inappropriately or irrationally prefer an option because it preserves the status quo.

Let's start with a general remark. Cutting oneself off the external world for limited periods of time can be – and often is – desirable. Watching television, playing video games, reading novels or going on holidays are ways to temporarily enter some form of experience machine – metaphorically speaking. In contrast to Nozick's experience machine, there are two crucial advantages to watching television or going on holidays, however. First, we remain alert to the external world to some degree. We don't have to put all our trust in some “superduper neuroscientists” and a supposedly perfectly functioning machine. Second, we remain agents capable of promoting our own happiness and the happiness of others.²⁹

If these conditions are not met, we consider activities of shutting oneself off the external world to be harmful because they put us in danger of ending up with less overall happiness due to long-term consequences, risks and uncertainties. That's why we consider it bad to take heroin. Given our own epistemological limitations, it rarely makes sense to put everything on the same card and completely trust somebody or something with our happiness and the happiness of those we feel responsible for. In a world as complex as our own, not even the hedonist should enter the experience machine when invited by “superduper neuroscientists”.

The hedonist cares about how the world really is. But not because there is some intrinsic value in authenticity or truth, but because it makes it considerably more likely to promote happiness when one has a firm grasp of reality. We shouldn't lose control and autonomy over our life because they are highly instrumental to our happiness in the actual world. The intuition invoked by the *Experience Machine* may thus be

debunked by pointing to its likely evolutionary causes, as Crisp [13, p. 122] suggests:

Valuing honesty, transparency, genuineness, and so on, has a clear pay-off: it fends off deception, and thereby assists understanding of the world, which itself issues in a clear evolutionary advantage.

The intuition that connectedness to reality matters is based both on our biology and experience. That is the first problem; it is extremely hard to imagine that there is such a perfect machine and such trustworthy neuroscientists that it would be more likely to be happier without our conscious control.

The second problem concerns our responsibility for other people. We naturally wonder what will happen to our loved ones and everybody else after entering the machine. Again, it can simply be stipulated that they are cared for. But what would such a possible world look like? If we try to imagine a world where everyone – every sentient being – is being cared for by trustworthy, benevolent and ingenious – in fact, nearly omniscient – neuroscientists who have created a perfect machine that will never break down and can generate happiness for billions of sentient beings, it becomes rather unclear what our intuitions tell us precisely. Such a world would be *very* different from the actual one.

And this is the third problem. When relying on our intuitive judgments, particularly in an underspecified and out-of-the-ordinary environment, we are likely to process irrelevant factors due to unconscious bias, as [54], [63], [38], [60], and [40] have pointed out. There is evidence that many people’s anti-hedonist intuitions about Nozick’s experience machine arise from status quo bias, for instance.³⁰ As Weijers [58, p. 101] puts it, “we have mislabelled what we are familiar with as ‘reality’”. Variants of the experience machine like the vignettes by de Brigard [7] or the *Trip to Reality* by Weijers [60] may show how intuitions change once we control for this bias.

These variants reverse the direction; instead of leaving reality and entering the machine, we have the choice to leave the machine and enter reality. Such reversal tests are to ensure we don’t mistake familiarity with reality. Consider Weijers’ *Stranger No Status Quo* thought experiment:

A stranger, named Boris, has just found out that he has been regularly switched between a real life and a life of machine-generated experiences (without ever being aware of the switches); 50% of his life has been spent in an Experience

Machine and 50% in reality. Nearly all of Boris' most enjoyable experiences occurred while he was in an Experience Machine and nearly all of his least enjoyable experiences occurred while he was in reality. Boris now has to decide between living the rest of his life in an Experience Machine or in reality (no more switching). [61, p. 525]

The question now is what would be best to do for Boris himself – to choose a life in the experience machine or in reality. We are asked to ignore “how Boris’ family, friends, any other dependents, and society in general might be affected, and assuming that Experience Machines always work perfectly” [61, p. 526]. While the *Stranger No Status Quo* avoids most problems related to status quo bias, there are still worries left concerning the reliability of our intuitions. Can we really just ignore our social connections to other people and the possibilities of the machine’s malfunction? If there is any uncertainty, it is difficult to tell how happy we would be in the long run for any decision.³¹ Also, we might simply be too curious for Boris to try his real life – anticipating future pleasure in reality from finding out.³²

Another way to adjust the *Experience Machine* may be to set it up in the past. Consider this alternative thought experiment:

Stranger in the Past

Imagine an experience machine which has been running smoothly for many centuries with many people having been plugged in. Two hundred years ago also Boris chose to plug in. We know that Boris spent 90 years in the machine until he died and that he has had a gloriously happy life there – one he wouldn’t have had outside the machine. Was his life in the machine better for him than the one he could have had outside?

Most problems of Nozick’s *Experience Machine*, Weijer’s *Trip to Reality* and *Stranger No Status Quo* are not present. Arguably, there is no status quo bias and no doubt about the functionality of the machine. Perhaps, more people would now say that Boris made the right choice. But there might still be people who have strong intuitions against Boris’s plugging in. What could the hedonist tell those who have the intuition that Boris made a huge mistake?

Let’s spell out the details to assess the soundness of the inference to the best explanation in premise (P^e) and the reliability of our intuitions in premise (P^r). How would a world look like in which such a machine

has worked perfectly for such a long time? What could Boris have done outside of the machine in such a world?

There is absolutely no way to tell because the existence of such a machine is so outside of our experienced universe that we cannot draw any reliable inference to the nature of such a world. The only thing that we know for sure is that such a world would be *extremely* different from our actual world – just as in the original *Experience Machine* thought experiment. Thus, even if no bias can be clearly identified, there is still reason to doubt our initial reactions to the thought experiment.

There is neither ground to back up the inference to the best explanation in premise (P^e) nor reason to believe that people can just rely on their intuitions (P^r). If we simply take the thought experiment's description at face value, we silently introduce irrelevant factors to the possible world. If we realistically spell out the details, we have to imagine a world which is so unlike our actual one that our intuitions become highly unreliable.

All three problems (the functionality of the machine, the role of other people and possible biases) concern the fact that we cannot easily disregard knowledge about us and the world when entertaining thought experiments. Even if we are explicitly told that some aspects are irrelevant, we anyway tend to take them into account in some way. In any even remotely realistic scenario, nobody – not even the most passionate hedonist – enters the experience machine. That is, if there remains some tiny little bit of doubt about the neuroscientists' or the machine's trustworthiness, not even the hedonist would enter it.

The Deceived Businessman Revisited

Kagan's *Deceived Businessman* is far more realistic, but faces essentially the same problems as Nozick's *Experience Machine*. The bias in the *Deceived Businessman* is (again) the following: We cannot easily discount information that we have.³³

For the thought experiment to work, we need to mentally separate the phenomenon of being deceived from its usual implications. We know that deceived people are generally worse off. So in any world remotely similar to the actual one, we have good reason to feel sorry for them. We cannot put ourselves entirely in the businessman's shoes, since that would mean to *unknow* something that we know.

The phenomenon is similar to cases in which people feel sorry for dead people when after their death it is found out that they have been cheated

on. It is sometimes rational to feel sorry if we have reason to suspect that the person cheated on had doubts and consequently suffered from jealousy or other worries. But even if we don't, we usually feel sorry.

A straightforward explanation lies in our inability to disregard information we have. In the *Experience Machine* we cannot easily discount the possibility that the neuroscientists are not as trustworthy as we are told. In Kagan's thought experiment we cannot easily discount the possibility that the businessman did somehow know or at least feel that he was being deceived – and, as a consequence, that there had been consequences to him in matter of experienced suffering.

But even if we could discount this information, we may say: “Good for him to have had such a life, but how terrible for his wife, children, friends and colleagues!” Again, if we want to imagine the scenario in some detail, we have to take into account the other people mentioned in the experiment. It will presumably be dreadful for them to live with such a delusional and self-opinionated man. It would be better for them not to have such a person around. Rarely anybody wants to be so deceived that they cannot recognize whether their fellow sentient beings are in pleasure or pain.

An additional aspect, which is relevant in the *Deceived Businessman*, is that we tend to evaluate from the point of view of the *remembering self*.³⁴ That is, we re-evaluate experienced happiness or suffering based on the knowledge that we have at the moment of remembering. Memories of one's own happiness have been shown to be highly unreliable.³⁵ There are usually stark differences between the *remembering self* and the *experiencing self*. In hindsight, we “remember” our holidays to be spectacular because we blend out the arduous flight, the mosquitoes and the arguments with the kids. Instead, we focus on the two minutes of tranquility during the sunset. That's fine because the process of remembering is itself an experience which can be more or less pleasurable. And, often, there is no reason to dwell on bad memories, even though in fact we have experienced less pleasure than we believe.

Sometimes, however, it is the other way around. We feel bad when we imagine that the businessman has been deceived because we imagine his happiness not as it was when he experienced it, but as it is remembered in our imagination – with the knowledge that it was all a lie. We know this phenomenon when people re-evaluate their romantic relationships once they have painfully ended. Suddenly, many happy moments become sad – not because they were, but because we put them in context with the new narrative of our life which includes the painful end of the

relationship.

Certainly, it would be painful for the businessman to wake up from his illusion at the end of his life and realize that he had been deceived during all that time. This is – by stipulation – excluded from the experiment. But this mental picture is at least part of what gives us the intuition that his life must have been worse than the one in which he was not deceived. We don't want to be deceived because – as a rule of thumb – deception leads to suffering. This rule is deeply rooted in our brains (both genetically and socially). If we cannot discount it, however, premise (P^e) is not warranted because we can easily explain our intuition in terms of pleasure and pain. And, even if people explicitly believe that the deceived businessman's life is worse, their conviction will not be formed in a reliable way. They subconsciously and systematically introduce irrelevant factors such as their feelings about deceived persons, the impact of deception on other people or hindsight evaluation. Then, premise (P^r) is not warranted.

3.2 *Intuitions and Values*

We can now draw a general lesson. On the one hand, if we take the thought experiment's description at face value and trust our intuitive impulse by imagining whatever comes to our mind, the argument poses no harm to hedonism because the initial reactions are easily explained; for any of the four arguments, there are several other plausible explanations why people show reaction R than their having anti-hedonist normative convictions.

On the other hand, if we interpret the experiment realistically and explicate the sensible ways to fill in the details, it poses no harm to hedonism either because the intuitions are not anti-hedonist anymore. In anything close to the real world the hedonist would never trade her life for that of the pig, choose the ugly world, enter the experience machine, or consider the life of the deceived businessman equally good. Only in a thought experiment that one entertains *faithfully and without questions asked* must the hedonist trade her life for the pig's one, be indifferent about choosing the beautiful or ugly world, enter the experience machine or consider the life of the deceived businessman equally good. But even the tiniest little bit of doubt makes it rational to do what everyone would do. In the real world, there is always a tiny bit of doubt. In the real world, we should steer clear from these choices because they cannot be expected to optimally promote happiness.

It is not a trivial task to fully explicate what a thought experiment entails. As Dennett puts it with respect to John Searle's *Chinese Room* thought experiment, intuition pumps discourage "us from attending to [all the supposedly irrelevant details]. But if we are to do a good job imagining the case, we are not only entitled but obliged to imagine [the scenario in a detailed way]. But then it is no longer obvious" that the life of the pig is less good, that the beautiful world is more valuable, that the experience machine should not be entered and that the deceived businessman has had a worse life.³⁶ Is the best explanation for the initial intuition really that there are qualitatively different pleasures or intrinsic values?

As a matter of fact, people value all sorts of things. People give value to paintings and cars, to justice and honor, to friendship and friends, to money and health. How we decide what to value depends on our intuitions and preferences. And, our intuitions and preferences depend on our biological constitution and our convictions about us and the world.

Our intuitions are shaped by the way we live and the way our ancestors used to live.³⁷ In situations which are part of or at least sufficiently similar to our and their lives this is fine and our intuitions can be expected to provide good prima facie evidence – but this evidence can be thwarted by unexpected or unusual conditions – which give rise to structural bias.³⁸ This precondition makes it a rather weak indicator of how we should evaluate situations that are very *unlike* our usual life. The best explanation for our initial reactions R may not be that we have anti-hedonist convictions. For all four thought experiments, there are multiple alternative explanations that are at least equally plausible. Hence, premise (P^e) is not warranted without further argument.

Moreover, our moral convictions in the thought experiments don't show us what *actually* is valuable. The four classical arguments against hedonism (as they are usually presented) presuppose that something's seeming valuable in the thought experiments is conclusive evidence that it is valuable.³⁹ It is doubtful, however, whether we are warranted in presupposing this and assuming premise (P^r). But this premise is crucial, since hedonism – properly stated – entails that a person's well-being *does* depend on the amount of happiness only. More to the point, it does *not* entail that people (say that they) intrinsically value happiness only. The hedonist estimates the overall happiness resulting from her options to decide which one to choose – even if this might be something that goes against her initial intuitive reaction. The intuitive reactions themselves are at best prima facie evidence for her moral judgments.

Thus, both premise (P^c) and, especially, premise (P^r) are in dire need to be justified independently. Neither of them is warranted as it stands and can be assumed without high theoretical costs. In contrast, giving them up doesn't have any particularly bad consequences for the quantitative hedonist.

Summary

The classical arguments made against quantitative hedonism are systematically flawed. The *Philosophy of Swine*, *Heap of Filth*, *Experience Machine* and *Deceived Businessman* are intuition pumps that prime us for mostly irrelevant aspects in out-of-the-ordinary contexts, while obscuring the causes of our intuitions. Once the salient aspects are identified as irrelevant and the causes of our intuitions are uncovered, there is no reason to assume that non-hedonistic values such as beauty, autonomy or truth cannot be properly accounted for as instrumental within quantitative hedonism. The thought experiments invoke intuitions that are irrevocably influenced by our knowledge about the real world in which there contingently are strong relations between autonomy, self-determination, beauty, truth, reality, etc. and happiness. Quantitative hedonism is thus committed to certain empirical claims about how we form our intuitions about values and the reliability of these intuitions – but it is not proven wrong by the arguments classically levelled against it.

Acknowledgements

I would like to thank my anonymous referees, the participants of the philosophical research seminar at Karlsruhe Institute of Technology (KIT) as well as those attending ISUS 2018 for helpful comments to earlier drafts of this paper and, in particular, Gregor Betz, Roger Crisp, Christian Seidel, Michael Schefczyk and Alexandra Zinke for their critical feedback.

Notes

- 1 Quantitative hedonism has most prominently been put forward by Jeremy Bentham and David Hume and is also sometimes called traditional hedonism.
- 2 In this paper, I use “hedonism” to refer to quantitative hedonism if it doesn't lead to confusion. I want to contrast it not only with non-hedonistic theories of

- well-being, but also with modern versions of hedonism such as the attitudinal hedonism of Fred Feldman [20,21]. Please also note that I use “hedonism” in a strictly normative sense that includes prudential or value hedonism that may provide the basis for hedonistic utilitarianism. Here, I am not interested in hedonism as a psychological theory of what people actually aim for in life.
- 3 There are many philosophers issuing similar verdicts. See, for instance, [39], [30], [49], [54], or [8].
 - 4 See, for instance, [36], [12], [56], [26], or [61].
 - 5 For example, Nozick did not originally use his Experience Machine thought experiment to argue against prudential hedonism. Instead, it was only adopted for this purposes later on (including by himself in [46]).
 - 6 See [10, p. 65]; the quotes are from [42].
 - 7 I don’t try to provide a historically accurate interpretation. This is one of several possible interpretations of this passage in Mill’s Utilitarianism. Some would prefer to understand it simply as polemics or an illustration of another – more complex – argument by Mill. These are viable interpretations, too. What I claim is merely that people have understood his argument thus and used it to attack hedonism.
 - 8 I use “happiness” here to denote the positive difference between pleasure and displeasure; I don’t mean eudaimonia or anything loaded with values other than experienced pleasure.
 - 9 Mill thus abandoned quantitative hedonism and replaced it with qualitative hedonism. While Acton [1, p. 541] calls quantitative hedonism “shocking but consistent” his verdict on Mill’s qualitative hedonism is even less favorable – calling it inconsistent. Some have tried to make sense of it as a version of quantitative hedonism, while others, such as Tännsjö [56, p. 97], consider it outside the “hedonist camp altogether”.
 - 10 Cf., e.g., [28].
 - 11 Cf. [12, p. 620] or [2, p. 257].
 - 12 Cf. [46, pp. 104ff.] for a slightly different version of the Experience Machine.
 - 13 This reconstruction resembles the abductive reconstruction by [62].
 - 14 This distinction is also drawn when hedonism is attacked by arguments based on the intuition that pleasures resulting from behavior as bad as sadistic torture are “false”. The existence of such “perverted” pleasures purportedly shows that well-being cannot depend solely on the amount of experienced pleasure. Cf. [24, p. 56], for the so called “evil pleasures objection” to hedonism.
 - 15 Cf. [31, pp. 43ff.].
 - 16 In each such instance, a theory or generalization is taken to be undermined by an intuition regarding a particular hypothetical case. Cf. [41].
 - 17 While Weijers and Schouten [62] have identified nearly the same argumentative structure with respect to the Experience Machine, their reconstruction is limited to this particular thought experiment and it is evaluated quite differently by the authors.
 - 18 Cf. [53].
 - 19 Dennett [17, p. 12] said: “Intuition pumps are cunningly designed to focus the reader’s attention on ‘the important’ features, and to deflect the reader from

- bogging down in hard-to-follow details. There is nothing wrong with this in principle. Indeed one of philosophy’s highest callings is finding ways of helping people see the forest and not just the trees.” Cf. also [16] or [37].
- 20 Intuitive judgments based on thought experiments seem likely to mislead. Cf. [14], [57], [23], [55], [37], [44], [63], [60]. Cf. [11] for an up-to-date discussion of the reliability of intuitive judgments in philosophical thought experiments.
- 21 Cf. Weijers [60, p. 22], who argues that “most philosophical thought experiments stipulate features that are so unrealistic that we have not experienced anything like them or we have experienced the very opposite of them. When these clashes occur, our intuitions are likely to be based on information that is not just irrelevant, but contrary to the point of the experiment itself. It is these features of intuitive cognition that enable structural biases to affect our judgments about thought experiments in ways that our deliberative judgment is not usually affected.”
- 22 But even if we successfully lift our cognitive reaction from the unconscious to deliberate thought, we are often stuck with post hoc rationalizations that are biased in other ways. Cf. section 3.2. For a detailed account, see also [63] or [60].
- 23 While it can be argued that the thought experiments’ authors themselves did not intend to infer anything from our intuitive reactions, most of the literature has (implicitly or explicitly) been taking them to do so. Compare [3], replying to the claim that philosophers do generally not rely on intuitions, for example - a claim put forward by [9] and [18].
- 24 There is good experimental philosophy on the Experience Machine: See, for instance, [47], [7], or [60]. There are some doubts on the reliability of such empirical surveys in experimental philosophy, though. The results should thus be interpreted with care. Cf. [51] for a (problematic) critique of [7]. For a (quite convincing) reply, see [59].
- 25 See [63] or [60]. Cf. also section 3.2.
- 26 Status quo bias is a type of cognitive bias in which status quo is taken as a reference point and any change from it is perceived as a loss. Cf. [48] or [34].
- 27 Also, there is a chance that becoming the happy pig is perceived as loosing the human way, evaluating the choice as one between incurring costs (becoming the pig) or not incurring costs (remaining the human being). Cf. [31].
- 28 Cf. Crisp’s remarks on the second edition of *Principia Ethica* in which Moore himself admitted that the book “is full of mistakes and confusions” [12, p. 620].
- 29 Cf. [26, pp. 342f.] for this point.
- 30 Cf. [5].
- 31 See also Silverstein [49], who argues that preferring not to enter the machine nearly always leads to more happiness in the long run.
- 32 It may be more telling how most people react to the skeptical Brain in the Vat scenario discussed in epistemology. According to Tännsjö [56, p. 95], “the standard reaction [...] is this: ‘So you mean that there is no way for us of telling whether we are brains in a vat or not? Well, in that case, who cares?’”
- 33 There are, of course, other responses in the literature to the Deceived Businessman. For instance, [19] tries to answer the challenge with his account of (Veridical Intrinsic) Attitudinal Hedonism. I want to examine, however, how the

unaltered version of quantitative hedonism fares with respect to Kagan's thought experiment.

34 See [35] for a differentiation of the remembering self and the experiencing self. Cf. also [33, pp. 311ff.].

35 See, e.g., [32].

36 See [15, p. 438].

37 See [50] for a discussion of the role of evolution with respect to moral intuitions.

38 Cf. Ichikawa [27, p. 108], who claims that there is no reason why the onus should be on the "philosophical theorist to explain away misleading intuitions. After all, we (almost) all agree that some intuitions are mistaken".

39 Cf. [26, p. 345].

David Lanius
Karlsruhe Institute of Technology (KIT)
Institute of Philosophy
Douglasstr. 24
76133 Karlsruhe, Germany
<david.lanius@kit.edu>

References

- [1] Harry Burrows Acton. „Animal Pleasures“. In: *The Massachusetts Review* 2 (3) (1961), 541–548.
- [2] Alex Barber. „Hedonism and the Experience Machine“. In: *Philosophical Papers* 40 (2) (2011), 257–278.
- [3] John Bengson. „How Philosophers Use Intuition and ‘Intuition’“. In: *Philosophical Studies* 171 (3) (2014), 555–576.
- [4] Jeremy Bentham. *The Rationale of Reward*. In: London: Robert Hewart. 1830.
- [5] Nick Bostrom, and Toby Ord. „The Reversal Test: Eliminating Status Quo Bias in Applied Ethics“. In: *Ethics* 116 (4) (2006), 656–679.
- [6] Francis Herbert Bradley. *Ethical Studies*. In: Oxford: Oxford University Press. 1927.

- [7] Felipe de Brigard. „If You Like It, does It Matter If It’s Real?“. In: *Philosophical Psychology* 23 (1) (2010), 43–57.
- [8] Bengt Brülde. „Happiness Theories of the Good Life“. In: *Journal of Happiness Studies* 8 (2007), 15–49.
- [9] Herman Cappelen. *Philosophy Without Intuitions*. In: Oxford: Oxford University Press. 2014.
- [10] Thomas Carlyle. *On Heroes, Hero-Worship, and the Heroic in History*. In: Berkeley, CA: University of California Press. 1840/1993.
- [11] Daniel Cohnitz. „Thought Experiments in Current Metaphilosophical Debates,“ *The Routledge Companion to Thought Experiments*. Ed. by Michael T. Stuart, Yiftach J. H. Fehige and James Robert Brown. Abingdon/New York: Routledge. 2018, 406–424.
- [12] Roger Crisp. „Hedonism Reconsidered“. In: *Philosophy and Phenomenological Research* 73 (3) (2006), 619–645.
- [13] Roger Crisp. *Reasons and the Good*. In: Oxford: Clarendon. 2006.
- [14] Daniel C. Dennett. „The Milk of Human Intentionality“. In: *Behavioral and Brain Sciences* 3 (3) (1980), 428–430.
- [15] Daniel C. Dennett. *Consciousness Explained*. In: Boston: Little Brown. 1991.
- [16] Daniel C. Dennett. *Intuition Pumps: And Other Tools for Thinking*. In: New York, NY: Norton & Company. 2014.
- [17] Daniel C. Dennett. *Elbow Room: The Varieties of Free Will Worth Wanting*. In: Cambridge, MA: The MIT Press. 2015.
- [18] Max Deutsch. *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*. In: Cambridge, MA: The MIT Press. 2015.
- [19] Fred Feldman. „The good life: A defense of attitudinal hedonism“. In: *Philosophy and Phenomenological Research* 65 (3) (2002), 604–628.
- [20] Fred Feldman. *Pleasure and the Good Life: Concerning the Nature, Varieties and Plausibility of Hedonism*. In: Oxford: Clarendon Press. 2006.

-
- [21] Fred Feldman. *What is This Thing Called Happiness?* vol. 24. In: Oxford: Oxford University Press. 2010.
- [22] Thomas Hill Green. *Prolegomena to Ethics*. In: Oxford: Oxford University Press. 1906.
- [23] Jonathan Haidt. „The emotional dog and its rational tail: A social intuitionist approach to moral judgment“. In: *Psychological review* 108 (4) (2001), 814–834.
- [24] John Harsanyi. „Morality and the Theory of Rational Behaviour,“ *Utilitarianism and Beyond*. Ed. by Amartya Sen and Bernard Williams. Cambridge: Cambridge University Press. 1982, 39–62.
- [25] Michael Hauskeller. „No Philosophy for Swine: John Stuart Mill on the Quality of Pleasures“. In: *Utilitas* 23 (4) (2011), pp. 428–446.
- [26] Sharon Hewitt. „What do Our Intuitions about the Experience Machine Really Tell Us about Hedonism?“. In: *Philosophical Studies* 151 (3) (2010), pp. 331–349.
- [27] Jonathan Ichikawa. „Explaining Away Intuitions“. In: *Studia Philosophica Estonica* 2 (2) (2010), 94.
- [28] Oliver Johnson. „Aesthetic Objectivity and the Analogy with Ethics,“ *Philosophy and the Arts*. Ed. by Godfrey Vesey et al., Royal Institute of Philosophy lectures. London: Macmillan. 1973, pp. 165–181.
- [29] Shelly Kagan. „Me and My Life“. In: *Proceedings Of The Aristotelian Society* 94 (1994), pp. 309–324.
- [30] Shelly Kagan. *Normative Ethics*. In: Boulder, CO: Westview Press. 1998.
- [31] Daniel Kahneman, and Amos Tversky. „Choices, Values, and Frames“. In: *American Psychologist* 39 (4) (1984), pp. 341–350.
- [32] Daniel Kahneman. „Experienced Utility and Objective Happiness: A Moment-Based Approach,“ *Choices, Values, and Frames*. Ed. by Daniel Kahneman and Amos Tversky. Cambridge: Cambridge University Press. 2000, pp. 673–692.
- [33] Daniel Kahneman. *Thinking, Fast and Slow*. In: London: Penguin Books. 2012.

- [34] Daniel Kahneman, Jack L. Knetsch, and Richard H. Thaler. „Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias“. In: *The Journal of Economic Perspectives* 5 (1) (1991), 193–206.
- [35] Daniel Kahneman, and Jason Riis. „Living and Thinking about It: Two Perspectives on Life,“ *The Science of Well-Being*. Ed. by N. Baylis, Felicia A. Huppert and B. Keverne. Oxford: Oxford University Press. 2005, 285–301.
- [36] Jason Kawall. „The Experience Machine and Mental State Theories of Well-Being“. In: *The Journal of Value Inquiry* 33 (3) (1999), 381–387.
- [37] Joshua Knobe. „Experimental Philosophy“. In: *Philosophy Compass* 2 (1) (2007), 81–92.
- [38] Adam J. Kolber. „Mental Statism and the Experience Machine“. In: *Bard Journal of Social Sciences* 3 (2009).
- [39] Will Kymlicka. *Contemporary Political Philosophy: An Introduction*. In: Oxford: Oxford University Press. 1990.
- [40] Katarzyna de Lazari-Radek, and Peter Singer. *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. In: Oxford: Oxford University Press. 2014.
- [41] Edouard Machery. *Philosophy Within Its Proper Bounds*. In: New York: Oxford University Press. 2017.
- [42] John Stuart Mill. *Utilitarianism*. In: London: Parker, Son & Bourn, West Strand. 1863.
- [43] George Edward Moore. *Principia Ethica*. In: Cambridge: Cambridge University Press. 1903.
- [44] Thomas Nadelhoffer, and Eddy Nahmias. „The Past and Future of Experimental Philosophy“. In: *Philosophical Explorations* 10 (2) (2007), 123–149.
- [45] Robert Nozick. *Anarchy, State, and Utopia*. In: New York: Basic Books. 1974.
- [46] Robert Nozick. *The Examined Life: Philosophical Meditations*. In: New York: Simon and Schuster. 1989.

-
- [47] Richard Rowland. „Our Intuitions About the Experience Machine“. In: *Journal of Ethics and Social Philosophy* 12 (1) (2017), 110–117.
- [48] William Samuelson, and Richard Zeckhauser. „Status Quo Bias in Decision Making“. In: *Journal of Risk and Uncertainty* 1 (1988), 7–59.
- [49] Matthew Silverstein. „In Defense of Happiness“. In: *Social Theory and Practice* 26 (2) (2000), 279–300.
- [50] Peter Singer. „Ethics and Intuitions“. In: *The Journal of Ethics* 9 (3) (2005), 331–352.
- [51] Basil Smith. „Can We Test the Experience Machine?“. In: *Ethical Perspectives* 18 (1) (2011), 29–51.
- [52] David Sobel. „Varieties of Hedonism“. In: *Journal of Social Philosophy* 33 (2) (2002), 240–256.
- [53] Roy A. Sorensen. *Thought Experiments*. In: New York/Oxford: Oxford University Press. 1992.
- [54] Leonard Wayne Sumner. *Welfare, Happiness, and Ethics*. In: Oxford: Clarendon Press. 2003.
- [55] Cass Robert Sunstein. „Moral Heuristics“. In: *Behavioral and Brain Sciences* 28 (2005), 531–573.
- [56] Torbjörn Tännsjö. „Narrow Hedonism“. In: *Journal of Happiness Studies* 8 (1) (2007), 79–98.
- [57] Peter Unger. *Living High and Letting Die*. In: New York, NY: Oxford University Press. 1996.
- [58] Dan Weijers. „Reality Doesn’t Really Matter,“ *Inception and Philosophy*. Ed. by David Kyle Johnson. Hoboken, NJ: Wiley-Blackwell. 2011, 92–107.
- [59] Dan Weijers. „We Can Test the Experience Machine: Reply to Smith“. In: *Ethical Perspectives* 19 (2) (2012).
- [60] Dan Weijers. „Intuitive Biases in Judgments about Thought Experiments: The Experience Machine Revisited“. In: *Philosophical Writings* 4 (1) (2013), 17–31.

- [61] Dan Weijers. „Nozick’s Experience Machine is Dead, Long Live the Experience Machine!“. In: *Philosophical Psychology* 27 (4) (2014), 513–535.
- [62] Dan Weijers, and Vanessa Schouten. „An Assessment of Recent Responses to the Experience Machine Objection to Hedonism“. In: *The Journal of Value Inquiry* 47 (4) (2013), 461–482.
- [63] James Woodward, and John Allman. „Moral Intuition: Its Neural Substrates and Normative Significance“. In: *Journal of Physiology* 101 (4) (2007), 179–202.