Fanciful Examples

Ian Stoner and Jason Swartwood

> "Fact, fact, fact!" said the gentleman. And "Fact, fact, fact!" repeated
> Thomas Gradgrind…. "You must discard the word Fancy altogether."
> --Dickens, *Hard Times*

A broad range of philosophical arguments turn on described cases. Described cases come in a variety of forms: real-world examples, extensively documented by historians or journalists (*Lest Innocent Blood Be Shed*, *Bad Blood*); compressed or condensed real-world case studies (common in biomedical ethics); rich fictional examples, such as episodes and characters from great novels and classic films (*Middlemarch*, *Huck Finn*); and some arguments turn on *fanciful* examples: short, wholly invented scenarios or characters that range from unlikely to wildly imaginative.

Should those who acknowledge a role for described cases in philosophy avoid fanciful examples? Are examples that depend on flights of imagination, or heavy stipulation of unlikely scenarios, less useful than real-world, historical, journalistic, or rich literary examples? In our experience some students, on first contact with fanciful examples, respond with Gradgrind-esque indignation. But it isn't just students. A number of professional philosophers have articulated general objections to the use of fanciful examples. We argue that general objections to fancifulness fail; a careful look at the role described cases play within the method of wide reflective equilibrium reveals that all examples succeed or fail according to a straightforward set of criteria that is independent of fancifulness.

1. What are described cases for?

Perhaps the most common use of described cases in philosophy is as material from which to begin or advance the process of wide reflective equilibrium (WRE).[1] It is the value of fanciful examples employed within the method of WRE that we defend, here.[2]

Wide reflective equilibrium is an iterative process by which we seek to achieve coherence between pre-theoretical considered judgments about particular cases; the theoretical principles that systematize, explain, and justify those considered judgments; and other relevant background knowledge about ourselves and our world (Rawls 1951; Daniels 1979, 2013; Cath 2016). The method seeks to develop a robust fit (or equilibrium) between specific judgments and general principles by, first, looking for cases of poor fit---cases where principles and considered

1

judgments conflict. Cases of poor fit are resolved by modifying either the principles or the judgments in question, and the process begins again. The result is that the coherence and internal justification of our set of beliefs improves over time; and we can reasonably hope that such incremental improvements move us closer to the truth (McDowell 1998, 36; Hursthouse 1999, 165-166).

Described cases play a central role within this method by providing vivid considered judgements that can be used to test the plausibility of specific beliefs; they are a kind of custom-built tool for illuminating the ill-fittingness of a target belief. Two examples will suffice to illustrate the role described cases typically play in philosophical arguments employing the method of WRE.

Many people, at least at first blush, believe that farming animals for meat needs no moral justification. Animals are animals, after all, and their meat tastes good. Norcross' Fred's basement example targets the belief that eating factory-farmed meat is morally permissible (Norcross 2004). Norcross sketches the story of Fred, a man who can only enjoy the taste of chocolate if he first primes his tastebuds with an hormonal extract prepared from the glands of tortured puppies. So Fred tortures and slaughters puppies in order to enjoy the taste of chocolate. Norcross expects most readers will form the judgement that Fred's behavior is morally wrong. But this judgement about the described case is in tension with the target belief that consuming factory-farmed animals is morally permissible. This tension between the target belief and our judgment about the described case can be resolved in various ways. Perhaps we should conclude that Fred *isn't* doing anything wrong, and work to adjust our gut reactions accordingly. Or perhaps we could find some relevant difference between Fred's behavior and the behavior of meat eaters. Norcross, however, argues that the tension should be resolved by jettisoning the target belief that eating factory-farmed meat is morally permissible.

Thomson's famous violinist example (Thomson 1971, 48-9) follows the same pattern. Many people believe that if fetuses have a right to life, then that "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" (48). The famous violinist example targets this belief. Thomson imagines a story in which you are kidnapped by the Society of Music Lovers and connected to the bloodstream of a famous violinist. If you remain connected to him for nine months, he will recover and go about his life. If you disconnect, he will die. Thomson expects most readers will share a reaction to this case: the violinist does not have a right to use your body to sustain his life (though it would of course be kind of you to let him use you). This belief about the described case is in tension with the target belief that a fetal right to life entails a fetal right to use its mother's body. There are various ways one might resolve this tension. Perhaps we should revise our judgment about the described case. (Perhaps the violinist has a right to our kidneys after all.) Or perhaps we could identify a morally relevant difference between unplugging from the violinist and terminating a pregnancy. Thomson, however, argues that the tension should be resolved by jettisoning the target belief and accepting that it is not the case that a fetal right to life entails a right to be carried to term.

Described cases intended to function as persuasive devices within WRE generally follow this pattern. An author expects a described case to elicit a specific belief among many or most of the people in the intended audience. That response is designed to be in tension with a target belief, which puts the reader into a kind of disequilibrium. The author argues that the response to the example is more reliable than the target belief, and so equilibrium should be restored by revising or rejecting the target belief.[3]

2. The common features of good described cases

If this is a fair characterization of the core persuasive role described cases play within the method of wide reflective equilibrium, then *all* examples, whether drawn from journalism, classic literature, or flights of fancy, are subject to the same criteria of success. We identify three ways a described case can fail; fanciful examples are not specially prone to these failures.

2.1 Most people in the intended audience are able to form a definite belief in response to the described case.

Derek Parfit's "Fifth Earthquake" is a described case that probably fails to meet this criterion. Parfit intends this example to tell against a moral principle he paraphrases from Susan Wolf. Thus, the belief "Fifth Earthquake" targets is relatively complex: "If we harm people, without their consent, as a means of achieving some aim, we thereby treat these people merely as a means, in a way that is always to be regretted, and that, if other things are equal, makes our act wrong" (2013, Vol 2, 145). Against this principle, Parfit sketches a case:

> You and your child are trapped in slowly collapsing wreckage, which threatens both your lives. You could save your child's life by using Black's body as a shield, without Black's consent, in a way that would destroy one of her legs. You could also save your own life, by causing Black to lose her other leg. But you believe that this act would be wrong, since it is only the saving of a child that could justify imposing such an injury on someone else. Acting on this belief, you save your child's life by causing Black to lose one leg. (145)

Is it wrong to sacrifice your life to save Black's leg, while simultaneously sacrificing Black's other leg to save your child's life? We find ourselves helpless to answer. The example is simultaneously complicated and under-described. There are many possible outcomes, each desirable end constraining the others in the fashion of a rickety Jenga tower. And yet, for all that, we are also denied information that seems relevant to evaluating the case; Black's reasons for refusing to sacrifice her own legs, especially, might be important to know.

Parfit composed "Fifth Earthquake" to elicit from most people in his audience two beliefs which, taken together, make it impossible to accept Wolf's principle. The beliefs Parfit expects to elicit are 1) saving your child's life at the expense of Black's leg is morally wrong and 2) your chosen course of action does not use Black merely as a means. If *anyone* other than Parfit clearly and confidently forms that pair of beliefs in response to "Fifth Earthquake," we would be mildly surprised; if the example works for most people in Parfit's target audience, we would be shocked.[4] But it isn't the *fancifulness* of "Fifth Earthquake" that is the problem. The problem is its intricacy and under-description.

As evidence that it's perfectly possible to form a definite belief in response to a fanciful example consider a semi-original described case more fanciful than "Fifth Earthquake" (and one we'll revisit in section 3.3). This example targets the belief that moral standing is contingent on biological species membership.

> *Alien Snout-Thing Encounter.* A spaceship lands and a huge septapede with a snout waddles down the gangplank. It asks if you could point it toward the nearest library, because it wants to learn everything it can about human culture, in order that it may better

grasp the broader cultural context of the episodes of *Ally McBeal* that have recently arrived at its homeworld. Is it morally permissible to kill and eat the snout-thing to see if its tissues taste good?[5]

The fanciful nature of "Alien Snout-Thing Encounter" presents no difficulty at all.[6] Our imaginations are easily able to accommodate its appeals to imaginary septapedes, flying saucers, and 90s television, because the scenario is simple and, though quickly sketched, includes the key information we require to form a judgment. Given how hard it is to form definite judgments about many real-life scenarios, it is unlikely that fanciful examples are more at risk than real-life examples of leaving their intended audience unsure how to respond. In fact, it is presumably the relative ease of forming judgments about well-designed fanciful cases that leads some philosophers to prefer them to subtler, more complicated, real-world cases. No matter what, when fanciful examples fail to meet the first adequacy criterion, they fail as persuasive tools for the same reason many real-life examples fail: we in the audience are sometimes unable to form an initial judgment about some described cases.

2.2 The definite belief formed in response to the example is a belief one is willing to stand behind---it is, after a bit of reflection, a credible judgment.

There are at least two reasons we might hesitate to endorse an initial response to a described case. First, we might, on reflection, decide that our initial clarity was illusory---on further reflection, it is hard to know what to think about it. Second, we might, on reflection, decide that our initial response depends on recognized cognitive or implicit biases, and so is not the sort of response we should rely on. Whatever the cause, a low-credence judgement about a described case is of little value for WRE.

A described case whose initial clarity is probably illusory is Parfit's "Future Tuesday Indifference." With this example, Parfit intends to target a belief partially constitutive of Humeanism about practical reason: "if a desire or pattern of concern does not involve theoretical irrationality, it cannot be open to rational criticism" (1987, 125). Against this belief, Parfit offers a delightfully wild described case:

> A certain hedonist cares greatly about the quality of his future experiences. With one exception, he cares equally about all the parts of his future. The exception is that he has *Future-Tuesday-Indifference*. Throughout every Tuesday he cares in the normal way about what is happening to him. But he never cares about possible pains or pleasures on a *future* Tuesday. Thus he would choose a painful operation on the following Tuesday rather than a much less painful operation on the following Wednesday.... This indifference is a bare fact. When he is planning his future, it is simply true that he always prefers the prospect of great suffering on a Tuesday to the mildest pain on any other day. (1987, 123--124)

Parfit expects readers to form the belief, in response to the case, that this hedonist's preferences are irrational. "Future Tuesday Indifference" is well-designed to generate that response. But that initial response is one that, on reflection, readers should treat with a relatively low level of confidence. The problem is that the act of empathic projection that Parfit asks of his readers---to imaginatively occupy the perspective of a man with Future-Tuesday-Indifference---is much harder than it appears on the surface.

Sharon Street (2009) makes this case at length, and we here highlight but one aspect of her discussion. In "Future Tuesday Indifference," Parfit invites us to imagine an unfamiliar kind of temporal preference discontinuity. When the clock strikes midnight Tuesday morning, our hedonist's preferences instantly re-shape themselves from a set that is indifferent to Tuesday pain to one that includes aversion to Tuesday pain. Imagine how such a scenario must play out: the non-Tuesday time-slices of our hedonist will try to dump all future pain onto the Future-Tuesday time-slices, indifferent to the knowledge that those Future-Tuesday time-slices will suffer terribly and try to rebel against the plans of the non-Tuesday time-slices. Parfit stipulates that the man with Future-Tuesday-Indifference has no false beliefs about personal identity, but that stipulation is difficult to square with oscillating preferences that leave different time-slices struggling against each other. On reflection, the task Parfit sets for us is more challenging than it first appeared; it is downright hard to imagine what it would be like to prefer to schedule future pain in a Tuesday-indifferent way, while having no false beliefs about personal identity. *Is* our hedonist's preference set intrinsically irrational? Or are his preferences instead callous, even hostile, toward strangely alienated time-segments of his future self? Until we have an answer to that question, we should hesitate to assign the level of credence to our judgement about the case that is required to usefully employ it in WRE.

Parfit anticipates the objection that this example is unreliable because it is "so bizarre" (1987, 124). But its fancifulness is not the proximate cause of the trouble. The trouble with the example is that the empathic projection required to evaluate it is difficult. In this case it's true that the strangeness of the example is what accounts for the difficulty of the empathic projection. But a wide variety of factors, including real-world cultural differences, differences in social position, and differences in lived experience, can render a specific empathic projection difficult to carry out (Draguns 2007).[7] It is, for instance, difficult for most Americans today to imaginatively occupy the perspective of someone born into slavery (Coates 2011). A described case that required readers to imaginatively occupy that perspective would also be a case we should be cautious about relying on, not because it is fanciful---many real people really were born into slavery---but because it is hard to be confident that one has succeeded in the required imaginative task.

A different reason we might assign low credibility to a definite belief formed in response to a described case is if that case involves appeal to recognized cognitive or implicit biases. Fanciful examples are probably less vulnerable to this problem than real-world or rich literary examples, for the reason that fanciful examples can be (and usually are) tuned to avoid or exclude the kinds of details that are recognized as inviting biased responses. But fanciful examples are not immune. For example, we believe judgments formed in response to Jeff McMahan's "Aphrodisiac" should be assigned a low level of confidence.

McMahan targets a conjunction of two beliefs commonly held by disability rights activists. 1) Prenatal screening for the purpose of eliminating disabled fetuses is morally wrong and 2) Disability is a form of diversity; disabilities are mere differences that do not make a person intrinsically worse off. McMahan argues that people who hold this pair of beliefs "seem to be committed to accepting the permissibility of deliberately causing disabling prenatal injury, even for frivolous reasons" (2005b, 129), and sketches the following case in response:

> Suppose there is a drug that has a complex set of effects. It is an aphrodisiac that enhances a woman's pleasure during sexual intercourse. But it also increases fertility by inducing ovulation.... In addition, however, it has a very high probability of damaging the new egg in a way that will cause any child conceived through the fertilization of that egg

to be disabled. The disability caused by the drug is not so bad as to make life not worth living, but it is a disability that many potential parents seek to avoid through screening. Suppose that a woman takes this drug primarily to increase her pleasure---if it were not for this, she would not take it---but also with the thought that it may increase the probability of conception; for she wants to have a child. She is aware that the drug is likely to cause her to have a disabled child, but she is eager for pleasure and reflects that, while there would be disadvantages to having a disabled child, these might be compensated for by the special bonds that might be forged by the child's greater dependency. (2005a, 90)

McMahan expects that you'll judge that it is wrong for the woman to take the aphrodisiac, and that judgement is designed to be in tension with the implications of the conjunction of beliefs he's targeting.

     As Elizabeth Barnes (2014) argues, one option for the disabilities rights advocate is to accept McMahan's judgment about the case but argue that it doesn't succeed in undermining the target belief.[8] But another option is to question the credibility of the judgment that taking the aphrodisiac is wrong. Barnes notes that "Aphrodisiac" expects readers to endorse common social attitudes toward potential mothers, toward the value of women's sexual pleasure, and toward disabled children (108). Barnes herself seems unsure what to make of the case (for her, it fails the first adequacy criterion) (109). But even for readers who form a definite belief in response to it, "Aphrodisiac" is not a useful case because it so clearly has the potential to elicit well-known implicit (even explicit) biases. Even if we do form the judgment McMahan expects, we should hesitate to stand behind that judgment with the level confidence required to employ it in the process of WRE.

     "Aphrodisiac" is fanciful, but that isn't the problem with it. The problem is that a bit of reflection reveals that the judgment it elicits is difficult to disentangle from well-known biases. But this is not a special problem of fanciful cases: real-world and rich literary examples (much like the real world itself) sometimes invite known biases. The appropriate response to all such cases is to reduce the degree of confidence we attribute to our judgments about them, which in turn undermines their value as de-stabilizers in the process of WRE.[9]

2.3 The belief formed in response to the described case is genuinely in tension with the target belief.

     Finally, consider the pair of examples from James Rachels's (1975) bare-difference argument. His target is the belief (enshrined in the American Medical Association Code of Ethics) that active euthanasia (helping a person die) is always wrong but passive euthanasia (allowing a person to die) is sometimes morally permissible. Rachels suggests that people often adopt this position on the basis of the underlying belief that "killing [is], in itself, worse than letting die" (79). Rachels targets this latter belief (which we could call the intermediary target belief), which in turn casts doubt on the ultimate target belief. "Smith and Jones" comprises two parallel stories of men with malign intent:

In the first, Smith stands to gain a large inheritance if anything should happen to his six-year-old cousin. One evening while the child is taking his bath, Smith sneaks into the bathroom and drowns the child, and then arranges things so that it will look like an accident.

In the second, Jones also stands to gain if anything should happen to his six-year-old cousin. Like Smith, Jones sneaks in planning to drown the child in his bath. However, just as he enters the bathroom Jones sees the child slip and hit his head, and fall face down in the water. Jones is delighted; he stands by, ready to push the child's head back under if it is necessary, but it is not necessary. With only a little thrashing about, the child drowns all by himself, "accidentally," as Jones watches and does nothing.

Does "Smith and Jones" elicit a clear response? For us, yes: Smith and Jones act comparably badly. Do we stand behind that reaction on reflection? Yes, we do. The problem, here, is that the paired examples are not genuinely in tension with either Rachels's intermediary or ultimate target beliefs.

Consider first Rachels's attempt to use the examples to show that the intermediary target belief (that, all else being equal, killing is worse than letting die) is false. Rachels wants to show that because killing is not intrinsically worse than letting die in the Smith and Jones cases it is therefore not intrinsically worse than letting die in any other cases - including instances of active and passive euthanasia. But, as Shelly Kagan has argued, this inference is fallacious: it falsely assumes that "if variation in a given factor makes a difference *anywhere*, it makes a difference *everywhere*" (1988, 12). To see why it is problematic for Rachels to make this assumption, it helps to consider other cases of killing and letting die.

For instance, in a response to Rachels's argument Philippa Foot imagines two cases she calls Rescue I and Rescue II (Foot 2014, 352). In Rescue I, you are rushing in your jeep to save five people who will otherwise drown in the ocean tide, and to do so you have to let one other person (who is in peril elsewhere nearby) die. In Rescue II, you are rushing to save the five people as in the previous example, but to do so you find out you would have to drive over -- and kill -- one other person trapped on the road. Suppose the consequences of your actions are the same in both cases. The disaster that you are letting the one person in Rescue I die from is similar in its effects to crushing the hapless person under your tires in Rescue II, and in each case you do something that results in five being saved and one dying. Suppose also your motivations in the two cases are the same. It is not unreasonable for a person to judge that the difference between killing and letting die *is* intrinsically morally relevant here: when all else is equal, letting one die to save five in Rescue I is morally permissible but killing one to save five in Rescue II is morally wrong. And, there's no obvious reason that this judgment is inconsistent with our judgment that Smith and Jones's actions are morally equivalent: there's no reason we can't say *both* that killing is not intrinsically worse than letting die in the Smith and Jones case *and* that killing is intrinsically worse than letting die in Rescue I and II. In fact, Foot argues that a plausible account of rights to non-interference and rights to goods and services can explain why the difference between killing and letting die is intrinsically relevant in some cases but not others (354). Regardless of whether her explanation is the correct one, the upshot is that the mere fact that the difference between killing and letting die is not intrinsically relevant in the Smith and Jones cases does not imply that it is not intrinsically relevant in other cases (such as Rescue I and Rescue II).

Thus, Rachels's examples are not genuinely in tension with the intermediary target belief: they do not show that the difference between killing and letting die is *never* intrinsically morally relevant. And, if that's so, then they are not obviously in tension with his ultimate target belief either: they don't by themselves give us reason to reject the belief that active euthanasia is always wrong but passive euthanasia is sometimes permissible. Supporters of the AMA policy apparently believe that killing is worse than letting die in the context of doctors treating patients

with terminal illness. But, "Smith and Jones" could only put pressure on *that* belief if "Smith and Jones" somehow established that killing is *never*, in itself, worse than letting die. Kagan and Foot rightly insist that the judgment that killing is not intrinsically worse than letting die in "Smith and Jones" does not generalize to the proposition that killing is *never* intrinsically worse than letting die.[10] Thus, even a clear, credible judgment that killing is not intrinsically worse than letting die in "Smith and Jones" fails to light up any meaningful tension with Rachels's ultimate target belief (that killing is, in itself, worse than letting die in the context of euthanasia).

Here again, the problem is not that the example is fanciful. The problem is that the examples are not genuinely in tension with the target belief(s). (Though for people who do indeed come to the table with the belief that killing a person is in every case morally worse than letting that person die, Rachels's paired examples are an effective persuasive tool, whether or not they are fanciful.)

3. Misguided general objections

In order to be effective as persuasive tools within the method of wide reflective equilibrium, any described case must elicit a judgment that is clear, credible, and in tension with the target belief. In section two we argued that fanciful examples are at no disadvantage in satisfying these criteria; in fact, their popularity might be due in part to the fact that they can be built, from the ground up, to satisfy them. In section three we turn to general objections to the use of fanciful examples. We argue that complaints developed in terms of fancifulness tend to be, at bottom, complaints about confused or misguided use of described cases in general.

3.1 Fanciful examples make for bad practical guides.

Objection: Thus says the objector: we should avoid fanciful examples because when we try to import lessons from fantasyland to the real world, we are likely to err. Perhaps nowhere is the problem clearer and more consequential than in the use of ticking bomb scenarios to justify torture. When asked to imagine a fanciful scenario in which we know that torture is the only way to generate reliable testimony that will certainly save thousands of lives, many people react by forming the belief that, in such a scenario, torture is morally permissible. The use of fanciful ticking-bomb scenarios thus leads us morally astray when we import that conviction back to the real world, where bomb-plots are rarely of the ticking variety, suspects are difficult to identify, and torture doesn't generate reliable testimony (Luban 2005).

Henry Shue spells out the problem this way:

> Why are imaginary examples like ticking-bomb hypotheticals so badly misleading about how to plan for real cases? They mislead in two different ways that compound the error: idealization and abstraction. Idealization is the addition of positive features to an example in order to make the example better than reality, which lacks those features. Abstraction is the deletion of negative features of reality from an example in order to make the example still better than reality. Idealization adds sparkle, abstraction removes dirt. Together they make the hypothetical superior to reality and thereby a disastrously misleading analogy from which to derive conclusions about reality. (Shue 2005, 231)

It is considerations like these that led him to conclude, all the way back in the 1970s, that "artificial cases make bad ethics" (Shue 1978, 141). This is not a narrow point about ticking

bomb scenarios; the vast gulf between fanciful examples and the real world should raise "serious doubts about the wide range of work that approaches moral and political dilemmas by attempting to uncover moral intuitions about exotic cases of the kind never or rarely encountered in ordinary life" (Sunstein 2005, 541).

Reply: This objection identifies an abuse of a philosophical tool and concludes the tool is the problem.

It would of course be bad practice to base public policy on fanciful examples instead of the facts on the ground in the real world. If we must decide whether to increase the reimbursement rate in Medicare, we cannot settle the question by asking if an increase would be efficient in an idealized and abstracted world with a single non-profit hospital. It would similarly be bad practice to seek personal moral guidance from fanciful examples. If you want to know whether to blow the whistle on shenanigans in your workplace, you cannot read the answer off an idealized and abstracted world in which your boss is the only boss and she is eager to root out corruption.

Properly employed, ticking bomb scenarios light up a tension with the belief that "torture is always, in principle, wrong." Accepting this neither provides practical guidance for CIA agents, nor settles the public policy question. One can reject the absolutist principle in light of ticking bomb scenarios and still hold the belief that "torture is overwhelming likely to be, *in practice*, morally wrong." It is poor practical reasoning to import lessons from an idealized and abstracted world directly to the real world, but it is perfectly good philosophy to use an idealized and abstracted world to light up the boundaries of beliefs.

Examples of theft illustrate the same point. We have a rich literary example of a morally permissible act of theft in *Les Miserables*. (There are undoubtedly many real-world examples of morally permissible or morally required thefts, as well.) Such examples show that nearly everyone believes that theft is sometimes permissible. They certainly don't show that there's nothing wrong with theft. It would be a mistake to read *Les Miserable* and conclude that, because Jean Valjean was right to steal the bread he needed to feed his nieces and nephews, theft is generally permissible and should be legal. But the mistake is in the reasoning---the transfer of a norm from a rich literary case to a real-world situation it little resembles. Within the method of wide reflective equilibrium, Jean Valjean's case is useful for establishing the implausibility of the overbroad moral principle that "theft is always, in principle, wrong." It is hardly the fault of this rich literary example if, in the wrong hands, is can be abused as a confused defence of petty crime.

Shue himself is sensitive to the difference between the method of reflective equilibrium, which can usefully employ fanciful examples, and the direct importing of norms and practices from fantasyland to the real world. This is why, in 2005, he amended his slogan from "artificial cases make bad ethics" to "do not base any institutional preparations on imaginary cases" (2005, 233). This clarified form of the slogan is exactly right. To base policy directly on imaginary cases would be an abuse of imaginary cases. To endorse *that* claim concedes nothing about their value as a tool of philosophical introspection and persuasion (Davis 2005, Spino 2012, Beck and de Wijze 2015). Fanciful examples help us improve our moral beliefs by testing them against clear-cut cases. Those beliefs, *conjoined with the real facts of the real world,* are what should guide our decision-making in the real world.

3.2 Fanciful examples yield the wrong kind of data for theory.

Objection: Shoemaker characterizes the method of fanciful examples in this way: "Many of us deploy [thought experiments] as our version of the scientific method. They isolate some feature of our experience and evoke intuitions about it, and these revealed verdicts enable us to adjust relevant theories in light of what we find" (Shoemaker 2015).

When an author describes a fanciful case, the resulting "data" is supposed to consist in the reaction to the case. But this could only be *useful* data if the reaction is the characteristic response of reliable moral knowers pointing their moral sensing apparatus at a potentially interesting case. There are, of course, concerns about how we could identify reliable moral knowers.[11] But even if we had a focus group of steely-eyed, infallible moral knowers, pitching them fanciful cases would be a waste of time, because fantasyland is not what we want data about.

Note, by way of illustration, that fanciful examples in the domain of moral responsibility (Frankfurt cases, for instance) are designed to generate stable and clear reactions about the presence or absence of freedom. These clear responses are totally unlike our gut reaction to attributions of freedom in the real world, which are almost never clear. The "data" we get in response to Frankfurt cases is clear-cut, while the "data" we get from our reactions to questions of responsibility in the real world is usually murky. This mismatch should be a red flag indicating that, when we rely on cases of responsibility in fantasyland, we gather data about fantasyland, and we mislead ourselves about the nature of responsibility in the real world when we theorize from that data. Even if we could somehow identify reliable moral observers, they couldn't generate useful data by pointing their moral lenses at fantasyland. That would be the equivalent of using a well-calibrated electron microscope to examine cartoons of insects. Such a boondoggle could not teach us anything interesting about real insects.

Reply: There is a core use-case of examples that invites analogies to the scientific method: the described case as counterexample to a moral theory. Many people teach Thompson's "Transplant" case[12] in way that encourages this analogy. Total welfare utilitarianism is a theory of ethics (a hypothesis) that makes prescriptions (predictions) about what is right in particular cases, and Thomson's "Transplant" is a thought experiment (lab experiment) in which the theory clearly yields the wrong prescription (prediction). "Transplant," one might be tempted to say, provides us with new data that falsifies the hypothesis of total welfare utilitarianism.

Though it's fun to talk this way, we should keep in mind that the analogy is loose, at best. "Transplant" doesn't give us data in the scientific sense. There is not a hard fact of the universe that we observe through the apparatus of "Transplant." Rather, the described case, like all the others we've discussed, is intended to put pressure on a target belief. In this case, the target belief is theory-linked.

A freshman, entirely innocent of the study of utilitarianism, could walk into a classroom already committed to the principle that morality requires us to do everything we can to save the greatest possible number of lives. Thompson's "Transplant" puts pressure on that belief by illuminating a tension between it and the student's belief about the described case. When "Transplant" is offered to an avowed total-welfare act-utilitarian, it is intended to function in exactly the same way. In neither case is "Transplant" offered as a fact---a datapoint---the theory fails to prescribe. In both cases it is offered to supply the audience with a vivid belief about the case that is in tension with the (potentially theory-linked) target belief.

To speak in scientific terms is misleading; the role of described cases is to illuminate internal tensions within a set of beliefs, not to turn a lens on the external world. Thus, the fact that fanciful examples do something other than faithfully represent the external world is no

weakness. They can still help us work---individually and, through discussion, collectively---toward a more coherent and better justified set of beliefs.

3.3 Fanciful examples stipulate away moral expertise.

Objection: Moral expertise in the real world consists, in large part, in two abilities: the ability to tell which of the various values and interests at play in a given situation are the salient ones and the ability to generate creative solutions that reduce conflict between various interests and values. Call these abilities perceptiveness and resourcefulness.

Part of what makes fanciful cases fanciful is that they stipulate away all sorts of real-world messiness and complication. In doing so, they usually eliminate any role for perceptiveness and resourcefulness. In a trolley problem, the stripped-down sketch of the situation exhaustively catalogs (so the story goes) the morally relevant features of the situation. *Perceptiveness* can play no role in trolley-land. And the other feature of trolley problems---that the author stipulates two options as the only available courses of action---likewise banishes *resourcefulness* from trolley-land (O'Connor 2012).

Concerning some of Derek Parfit's more fanciful cases, Philip Kitcher writes, "Readers are pitched into a fantasy world, remote from reality, in which our natural reactions are sharply curtailed by authorial fiat" (Kitcher 2012). In such a world, in which what we recognize as real-world moral expertise is sidelined, and our experience of real-world moral tensions is distorted beyond recognition, we can learn nothing of value (Wolff 2015). In trolley-land, or any other tiny universe of fanciful examples, the kind of moral wisdom we respect and seek is useless.

Reply: Consider again the belief that biological personhood is morally decisive; creatures with human DNA are sacrosanct, and creatures without it are fair game. This is not a belief characteristic of a moral expert.[13] One of the virtues of "Alien Snout-Thing Encounter" is that it is useful as means of philosophical persuasion for an audience of moral *non*-experts. It can help anyone who genuinely engages with it to see that "human DNA is what makes you matter, end of story" is not a belief that can withstand reflective scrutiny.

That fanciful examples assume no moral expertise (and in most cases sideline it entirely) would matter a great deal if the beliefs of moral experts were important to their persuasive role. But that is not how described cases work. Described cases, whether fanciful or realistic, are intended to persuade people into changing their minds, not by showing them that their beliefs are in tension with the beliefs of an expert, but rather by showing them that their *own* avowed beliefs are in tension with their *own* beliefs about the described case. Improving our set of moral beliefs in this way is worth doing, even though this effort is distinct from the development of the virtues of perceptiveness and resourcefulness that are characteristic of moral experts in the real world.

4. Conclusion

When described cases are properly employed as tools of introspection and persuasion within the method of wide reflective equilibrium, their degree of fancifulness plays little role in determining their success. It is only when they are misemployed that fancifulness is a weakness. General dismissal of fanciful examples thus distracts attention from a project that is important and remains unfinished: a careful analysis of the typical features of philosophically useful described cases.[14]

References:

Barnes, Elizabeth. 2014. "Valuing Disability, Causing Disability." *Ethics* 125, no. 1:88-113.

Beck, Simon and Stephen de Wijze. 2015. "Interrogating the 'Ticking Bomb Scenario': Reassessing the Thought Experiment." *International Journal of Applied Philosophy* 29, no. 1:53-70. DOI: 10.5840/ijap201561741

Brandt, Richard. 1990. "The Science of Man and Wide Reflective Equilibrium." *Ethics* 100, no. 2:259-278.

Butler, Octavia. 2012. "Bloodchild." In *Bloodchild: And Other Stories*. Open Road Media.

Cath, Yuri. 2016. "Reflective Equilibrium." In *The Oxford Handbook of Philosophical Methodology*, edited by Cappelen, Gendler & Hawthorne, 213-230. Oxford, UK: Oxford University Press.

Chambers, Tod. 1999. *The Fiction of Bioethics: Cases As Literary Texts*. Psychology Press.

Coates, Ta-Nehisi. 2011. "A Muscular Empathy" blog post of December 14th, 2011. http://www.theatlantic.com/national/archive/2011/12/a-muscular-empathy/249984/

Cohen, Carl. 1986. "The Case for the Use of Animals in Biomedical Research." *The New England Journal of Medicine* 314:865-869.

Daniels, Norman. 1979. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76, no. 5:256-82.

Daniels, Norman. 2013. "Reflective Equilibrium." In *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2013/entries/reflective-equilibrium/>.

Davis, Michael. 2005. "The Moral Justification of Torture and other Cruel, Inhuman, or Degrading Treatment." *International Journal of Applied Philosophy* 19, no. 2:161--78. DOI: 10.5840/ijap200519215

Draguns, Juris. 2007. "Empathy across National, Cultural, and Social Barriers." *Baltic Journal of Psychology* 8, nos. 1&2:5-20.

Drwecki, Brian B., et al. 2011. "Reducing Racial Disparities in Pain Treatment: The Role of Empathy and Perspective-Taking." *Pain* 152, no. 5:1001–1006.

Foot, Philippa. 2014. "Killing and Letting Die." In *Disputed Moral Issues*, 3rd ed. Oxford University Press.

Gendler, Tamar Szabó. 2000. *Thought Experiment: On the Powers and Limits of Imaginary Cases*. Garland Publishing.

Gendler, Tamar Szabó. 2007. "Philosophical Thought Experiments, Intuitions, and Cognitive Equilibrium." *Midwest Studies in Philosophy* 31, no. 1:68-89.

Hursthouse, Rosalind. 1999. *On Virtue Ethics*. New York, NY: Oxford University Press.

Kagan, Shelly. 1988. "The Additive Fallacy." *Ethics* 99:5–31.

Kant, Immanuel. 2002. *Groundwork for the Metaphysics of Morals*, trans. Allen Wood. Yale University Press.

Kitcher, Philip. 2012. "The Lure of the Peak." *New Republic*, February 22. Available online: http://www.newrepublic.com/article/books/magazine/99529/on-what-matters-derek-parfit

Knobe, Joshua. 2003. "Intentional Action in Folk Psychology: An Experimental Investigation." *Philosophical Psychology* 16:309-324.

Luban, David. 2005. "Torture, American Style." *The Washington Post*, Sunday, Nov. 27th, 2005. Available online: http://www.washingtonpost.com/wp-dyn/content/article/2005/11/25/AR2005112501552.html

Marquis, Don. 1989. "Why Abortion Is Immoral." *The Journal of Philosophy* 86, no. 4:183-202.

Marquis, Don. 2001. "The Moral-Principle Objection to Stem Cell Research." *Metaphilosophy* 38, no. 2-3:190-206.

McDowell, John. 1998. *Mind, Value, and Reality*. Harvard University Press.

McMahan, Jeff. 2005a. "Causing Disabled People to Exist and Causing People to Be Disabled." *Ethics* 116, no. 1:77-99.

McMahan, Jeff. 2005b. "The Morality of Screening for Disability." *Reproductive Biomedicine Online* 10, Supplement 1:129-132. DOI:10.1016/S1472-6483(10)62221-3

Nichols, Shaun and Joshua Knobe. 2007. "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions." *Nous* 41:663-685.

O'Connor, James. 2012. "The Trolley Method of Moral Philosophy." In *Essays in Philosophy* 13:242-255.

Parfit, Derek. 1987. *Reasons and Persons*. Oxford University Press.

Parfit, Derek. 2013. *On What Matters.* Oxford University Press.

Rachels, James. 1990. *Created from Animals: The Moral Implications of Darwinism.* Oxford University Press.

Rachels, James. 1975. "Active and Passive Euthanasia." *The New England Journal of Medicine* 292, no. 2:78-80.

Rawls, John. 1951. "Outline of a Decision Procedure for Ethics." *The Philosophical Review,* 60, no. 2:177-97.

Sabin, Janice A., et al. 2009. "Physicians' Implicit and Explicit Attitudes About Race by MD Race, Ethnicity, and Gender." *Journal of Health Care for the Poor and Underserved* 20, no. 3:896–913.

Shoemaker, David. 2015. "Thought Experiments in Philosophy." Blog post of April 26th 2015. http://blog.oup.com/2015/04/thought-experiments-philosophy/

Shue, Henry. 1978. "Torture." *Philosophy & Public Affairs* 7, no. 2:124--143.

Shue, Henry. 2005. "Torture in Dreamland: Disposing of the Ticking Bomb." *Case W. Res. J. Int'l L. 37*:231-239.

Singer, Peter. 1972. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1, no. 3: 229-243.

Singer, Peter. 2005. "Ethics and Intuitions." *The Journal of Ethics* 9, no. 3:331--352.

Spino, Joseph. 2012. "Defusing Dangers of Imaginary Cases." *International Journal of Applied Philosophy* 26, no. 1:29-37.

Street, Sharon. 2009. "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters." *Philosophical Issues* 19, no. 1:273-298.

Sunstein, Cass. 2005. "Moral Heuristics." *Behavioral and Brain Sciences* 28, no. 4: 531-541.

Thomson, Judith Jarvis. 1971. "A Defense of Abortion." *Philosophy & Public Affairs,* 1, no. 1:47--66.

Thomson, Judith Jarvis. 1985. "The Trolley Problem." *The Yale Law Journal,* 94, no. 6:1395-1415.

Warren, Mary Anne. 1973. "On the Moral and Legal Status of Abortion." *The Monist,* 57, no. 1:43--61.

Wolff, Robert Paul. 2015. "Something Hard." Blog post of May 15th, 2015. http://robertpaulwolff.blogspot.com/2015/05/something-hard.html

Wyndham, John. 2001. *The Day of the Triffids.* Penguin UK.

1An alternative and unobjectionable use of described cases is as concrete illustrations of general claims, as when Kant illustrates the good will by means of the shopkeeper example (*Groundwork*, Ak 397) or Peter Singer illustrates his minimal principle of beneficence with the example of a child drowning in a shallow pond (1972, 231).

2The method of WRE is not entirely uncontroversial. Prominent critics of the method include (or would include) Immanuel Kant (*Groundwork*, Ak 408-409), Peter Singer (2005, 345), Richard Brandt (1990), and others. We won't address these general rejections of the method. Our target audience is people who accept WRE as a philosophical method, but are skeptical that *fanciful* examples can be *good* or *useful* examples within that method.

3Although Gendler focuses on the psychological mechanisms that underlie our tendency to react differently to different descriptions of the same situation, her characterization of the role of thought experiments in philosophical argument is compatible with the characterization of the role of described cases we offer here. Gendler: "Thought experiments recruit representational schemas that were otherwise inactive.... Thought experiments can be expected to produce responses to the target material that remain in disequilibrium with responses to the same material under alternative presentations.... When thought experiments succeed as devices of persuasion, it is because the evoked response becomes dominant" (Gendler 2007, 69).

4Parfit, of course, has different expectations of his audience. "Most of us would believe that, in saving your child's life by destroying one of Black's legs, you would be acting wrongly. This, I assume, would also be Wolf's view. But Wolf's Principle supports this view only if we can truly claim that you are treating Black merely as a means. And as I have said, that claim is false, since you are giving up your life for Black's sake" (146).

5We expect you believe it would be morally wrong to eat the snout-thing. That's why similar examples appear in Mary Anne Warren's and Don Marquis's abortion papers and James Rachels's work on duties to animals. See Warren 1973, Marquis 1989, 191 and Rachels 1990, 183. We further acknowledge our debt to episode twelve of *Futurama*, "When Aliens Attack," for first suggesting probable interstellar interest in *Ally McBeal*.

6Readers of speculative fiction already know that the problem with examples like Parfit's can't be their fancifulness. Authors of speculative fiction regularly spin elaborate flights of fancy that nevertheless elicit powerful reactions from readers. Two of our favorite examples are Wyndham's *Day of the Triffids* and Butler's "Bloodchild." These stories are stranger than most described cases in philosophy, and yet readers' imaginations are nimble enough that they have for decades had emotionally and morally intense reactions to them.

7In other cases, implicit biases prevent us from seeing when we're failing to take successfully another's perspective. For instance, evidence suggests that health care professionals' failures of empathy (likely mediated by implicit racial bias) result in worse pain care for African American patients (Drwecki et al. 2011), and physicians often self-report having lower levels of bias than they actually exhibit (Sabin et al. 2009).

8In this case, that would mean trying to show that the wrongness of taking the aphrodisiac doesn't imply that we should reject the belief that causing pre-natal disability is morally permissible. For instance, we could try to show that taking the aphrodisiac is wrong not simply because it causes disability, but rather because it does so for bad reasons.

9Some (e.g. Singer 2005) believe there is evidence that human judgment is so systematically biased in such a wide variety of ways that the method of described cases (including realistic cases) should be rejected entirely. This seems to us misguided because it overestimates the pervasiveness of the problem, it underestimates the potential for the same problem to affect methods that don't involve described cases, it overlooks the necessity of using described cases to specify and test abstract principles, and it unjustifiably assumes that our beliefs about abstract principles are, in general, somehow more reliable than our beliefs about described cases. But defending the method of wide reflective equilibrium is beyond the scope of this paper. For an overview of the unreliability-of-intuitions objection to WRE, and some possible responses to it, see Daniels 2013. See also Daniels 1979 and Rawls 1951.

10Kagan makes this point via his own pair of bare-difference cases. The target belief, ostensibly held by Rachels on the basis of cases like "Smith and Jones," is that there is *never* any intrinsic difference between killing and letting die. Consider "Ludwig and Sylvia":

> Ludwig sees Sylvia drowning, but since the rocks beneath the water would do extensive damage to his boat, he decides not to rescue her.
> Ludwig sees that his boat is about to hit Sylvia, but since avoiding her would mean steering into the rocks, which would do extensive damage to his boat, he decides not to change course. (1988, 7--8)

Kagan expects readers will judge that Ludwig's heinous behavior in the second case (killing) is worse than his horrible behavior in the first (letting die). If this judgment withstands critical reflection, then it is in tension with Kagan's target belief: the assumption (made by Rachels) that if a feature is intrinsically (ir)relevant in one case, then it is intrinsically (ir)relevant in all cases.

11Compare the case of observations in empirical science: the data generated by an electron microscope is only useful if we know the microscope is well-designed and functioning properly. We have ways to test the reliability of microscopes; we have no way to test the reliability of reader reactions to described cases. In fact, we know full well that different audiences, and different people, have different reactions to the same described case. This has been demonstrated repeatedly by experimental philosophers (e.g., Knobe 2003, Nichols and Knobe 2007). When people have different reactions to the same case, whose reaction should we pay attention to? Who is the reliable source of "data?"

This is not an objection to fanciful examples, but rather to the method of WRE, and so is beyond the scope of this paper.

12 This fanciful example certainly meets the first two adequacy criteria we suggest above:

> [I]magine yourself to be a surgeon, a truly great surgeon. Among other things you do, you transplant organs, and you are such a great surgeon that the organs you transplant always take. At the moment you have five patients who need organs. Two need one lung each, two need a kidney each, and the fifth needs a heart. If they do not get those organs today, they will all die; if you find organs for them today, you can transplant the organs and they will all live. But where to find the lungs, the kidneys, and the heart? The time is almost up when a report is brought to you that a young man who has just come into your clinic for his yearly check-up has exactly the right blood-type, and is in excellent health. Lo, you have a possible donor. All you need do is cut him up and distribute his parts among the five who need them. You ask, but he says, "Sorry. I deeply sympathize, but no." Would it be morally permissible for you to operate anyway? Everybody to whom I have put this second hypothetical case says, No, it would not be morally permissible for you to proceed. (Thomson 1985, 1396)

13 It is of course possible to find among philosophers defenders of some version of this belief (Cohen 1986).

14 For some useful work on described cases, see Gendler 2000 and Chambers 1999.