Paul Bloom - *Reason Restored*

Taken from *A Very Bad Wizard: Morality Behind the Curtain, 2nd Edition* by Tamler Sommers
In the first edition of *A Very Bad Wizard: Morality Behind the Curtain – Nine Conversations*, philosopher Tamler Sommers talked with an interdisciplinary group of the world’s leading researchers—from the fields of social psychology, moral philosophy, cognitive science, and primatology—all working on the same issue: the origins and workings of morality. Together, these nine interviews pulled back some of the curtain, not only on our moral lives but—through Sommers’ probing, entertaining, and well informed questions—on the way morality traditionally has been studied.

This Second Edition increases the subject matter, adding eight additional interviews and offering features that will make *A Very Bad Wizard* more useful in undergraduate classrooms. These features include structuring all chapters around sections and themes familiar in a course in ethics or moral psychology; providing follow-up podcasts for some of the interviews, which will delve into certain issues from the conversations in a more informal manner; including an expanded and annotated reading list with relevant primary sources at the end of each interview; presenting instructor and student resources online in a companion website.

The resulting new publication promises to synthesize and make accessible the latest interdisciplinary research to offer a brand new way to teach philosophical ethics and moral psychology.
Much of the 2500 year history of moral psychology has focused on two questions. The first is descriptive: how do we arrive at our moral judgments? The second is normative: how should we arrive at our moral judgments. And for almost all of the years, the dominant answer to both questions was the same: reason. Plato’s image of the chariot from the *Phaedrus* summed it up best. Our emotions and our sense of honor are the horses, and reason is the charioteer. That’s how it is and that’s how it should be.

In the new century, however, the tide began to turn. Jonathan Haidt summarized the shift in his article “The Emotional Dog and His Rational Tail.” Evidence in behavioral economics and social psychology suggested that our emotions were the primary factors driving our moral judgments. Reason was relegated to the role of “press secretary” (Haidt’s phrase)—offering post-hoc justifications for judgments we already reached through our emotional responses. And thanks in part to a renewed interest in virtue ethics, as well as the work of feminist philosophers defending an ‘ethics of care,’ philosophers began to take emotions seriously again as proper guides for moral judgment. Reason, many argued, focuses too much on abstract idealizations; the emotions are more sensitive to context, to the particulars of real-life moral questions and problems.

Not so fast, says Paul Bloom, the Brooks and Suzanne Ragen Professor of Psychology at Yale University. Bloom’s research has covered a wide range of topics, from our intuitive dualism (a belief that world contains both matter and immaterial souls) to the psychology of pleasure to baby morality, and much more. His new project sets its sights on this new paradigm. In two recent articles for The New Yorker and The Atlantic, Bloom argues that the pendulum has swung too far against reason on both the descriptive and normative questions. On the descriptive side, Bloom objects to the new view of human beings as “biochemical puppets,” driven by wholly irrational forces over which we have no control. While conceding that these influences exist, Bloom argues that they do not undermine the idea of a responsible agent, capable of rational deliberation. On the normative side, Bloom has taken a hard line stance against the idea of emotions as a guide to good policy. And he does not pick easy targets. Many researchers warn against the effects of emotions like anger, envy, and revenge feelings, but Bloom has taken on the most beloved moral emotion of them all: empathy.

Bloom is the author of four books, most recently *Just Babies: The Origins of Good and Evil*. He is a wonderfully clear and lively writer, appealing both to scholarly and popular audiences, often in the same work. (He is also a favorite guest on *The Very Bad Wizards* podcast.) And though Bloom may be empathy’s arch nemesis, he is also one of nicest and funniest people I have encountered in this profession. I interviewed him at his
beautiful home in New Haven, CT. Observing a Jewish tradition that dates back nearly as far as the history of moral psychology, we argued, debated, and then had take-out Chinese. – June 2014

1. A Pox on Empathy!

**TS:** You recently wrote a piece for The New Yorker that was critical of one of our more cherished emotions: empathy. First, how do you define empathy and how do you distinguish from sympathy?

**PB:** The terminology here is murder. People use these words in very different ways. For me, empathy is feeling what you think other people feel, putting yourself in their shoes. So, if you bang your knee and I wince, that's empathy. So empathy is feeling someone else's pain or pleasure. Sympathy, or compassion (and I use the terms interchangeably) is caring about other people, caring about their fates.

**TS:** So the way you define it, empathy is when I feel your pain. But it's a further step for me to care that you're hurt and want to help you.

**PB:** Exactly. We can best understand the distinction by seeing them at work and seeing how one doesn't entail the other. You can have empathy without sympathy. Suppose you see me on the street and I'm in terrible pain and you feel bad too. Well, that's empathy. It's possible that could lead you to say "I care about that guy" and then move to help me. But it's also possible that you could turn away - my pain is causing your pain, so you turn away and walk away. It's even possible you could give me a kick and get mad at me for causing your pain. That would be empathy without sympathy.

**TS:** How about sympathy without empathy?

**PB:** Well, I could really care for somebody and not put myself in their shoes. So, here's an example. This example isn't mine but I forget whose it is. You watch people you love make a fool of themselves and you feel terrible. But they don't know they're making a fool of themselves. They're telling some incredibly offensive joke, and you're feeling mortified for them. This can't be empathy, because they're feeling fine - they're great. But you love them, and you know what they're doing to themselves. You feel compassion, you care about them.

**TS:** Right, although if someone is making a fool out of themselves, then maybe what I'm feeling is the emotion they would have "if they only knew". The big hypothetical: if they knew how they looked or how people were looking at them, how embarrassed they would be. Because otherwise, why would I feel embarrassed? You're right that this doesn't match up very well with empathy, but it also doesn't match up well with the kind of cognitive, or rationally motivated judgment.

**PB:** I think you're right, and the conclusion from that is that sympathy can be a very
emotional thing. I was just at this conference this month and I talked to this monk Mattieu Ricard. He was such a cool guy.

TS: I read his book, the one he did with his Dad who's a famous atheist French philosopher. It was fantastic, inspiring.

PB: I'm probably the least spiritual person in the world, and I didn't think we'd get along. But, I'm giving my riff against empathy and he agrees with me, which stunned me. He says, "I don't want to nurture empathy. Empathy makes you tired, empathy makes you miserable. I want love." And in fact, he says that to be a good person, and to feel love toward others, don't experience what they're experiencing. Keep your distance, at some level, but feel love and compassion for them. And I find that powerful, and I agree with that. And he's a monk so...

TS: Aside from the fact that he's a monk, why do you agree?

PB: An example I give is actually from an experiment by Dan Batson, psychology's most important altruism researcher. He has this experiment where there's a girl and she's waiting in line for a medical intervention. Batson gets people to take her perspective, triggering empathy, which triggers compassion, which makes them very likely to want to move her up the list. And in fact as we're speaking there's a news story about a little girl who is waiting to get a lung but she can't because of the rules. And people are petitioning. Batson points out correctly that it's a very compassionate response. But it's also not moral.

TS: What do you have to assume about morality to say that?

PB: That morality entails some sort of fairness and impartiality.

TS: You think that's inherent and necessary to morality?

PB: I think it's at the core of morality.

TS: One objection I had to your critique of empathy is that it seemed like you were holding empathy to an impossible standard. It was as if someone wrote an article called "The Case Against Blueberries" or "The Case Against Kale" and they say: "look, yes, blueberries are nutritious, really delicious, but they can't cure cancer. Not only that, you can eat too many blueberries (or too much kale) and get sick. So, blueberries and kale can't solve all of our health problems.

Well, yeah, but no one ever said they could. With empathy, it seems like we have something that's morally healthy for the most part, and an important foundation for morality. But instead of dwelling on that aspect of empathy you dwelled on its shortcomings. If I'm Empathy responding to you I'd say: "Well hey – I never pretended I was going to solve all moral problems and lead to all the correct moral judgments."
**PB:** I’m a fan of empathy for certain things. When it comes to intimate relations, I think, it is indispensable. I would not want a child, a parent, or a wife without empathy. Or a friend without empathy who couldn’t make an effort to feel what I feel - not just to understand it, but to feel what I feel. You know, if I’m really sad, and I want somebody close to me, I want some of that sadness to move on to them. I think that’s part of a relationship. And if I’m happy, I want my happiness to move on to that person. And to find somebody who lacked that would be a strange relationship, even putting aside the cognitive inabilities that would go with not feeling that, there’s also just a lack of personal contact.

So, I’m down with empathy for that. But I think empathy is horrible when it comes to public policy. The reason why this isn’t quite analogous to kale and blueberries is that there are people, including our president, who champion empathy whenever it comes to public policy. There are people like Jeremy Rifkin who say that the solution to all our global problems is an empathetic community. Empathy will solve the world’s problems. I think they’re just flat out mistaken, and I think this is a serious mistake because it’s not just that empathy is merely only part of the story. Actually it often leads us astray. In that article, I give some examples where empathy causes waste, suffering.

**TS:** So is it fair to say that you’re against the overblown rhetoric about empathy rather than empathy itself?

**PB:** More than that, I’m against using empathy as a way to make policy decisions. And I think that’s done a lot.

**TS:** But when you’re trying to come up with policies about poverty, don’t you want to have some sort of empathetic understanding of what it’s like to be poor, or to go to an inner city school? Or think of our drone policies. When we decide about that, don’t you want to consider how it feels to have drones circling in the sky, hovering over you round the clock? That’s something you would want, not merely a cold blooded analysis of the extent that drones will accomplish certain goals.

**PB:** There’s something deep in what you’re saying that I haven’t emphasized enough, which is that to make any interesting policy decision you want the information about people’s experiences. It’s incredibly important. So, before launching drones, find out how people respond to drones, how people respond to this and that. And that requires being able to look inside people’s heads in a sense, feel what they feel, and there I think understanding is relevant. Some people call this “cognitive empathy” and that’s fine. This isn’t the sense of empathy that I’m arguing against.

One way to express my beef with empathy is that for any interesting policy decision, there’s going to be factors pulling you in both directions. And that’s just the way any hard decision goes. And, empathy can be used to tip the balance either way.
TS: You mean that both sides can use it?

PB: Right. You want drones? Tell me a story about the victims of terrorism. You don’t want drones? Tell me a story about somebody who had his family blown up by a drone - it could go either way. You want gun control? Focus on the victims of a gun accident, gun crime. You’re against gun control? Tell me about a defenseless victim.

TS: So you’re against not empathy, but selectively applied empathy or empathy as the only guide to policy, right?

PB: No, it’s stronger than that. Empathy is by definition selectively applied. Not even randomly. We feel empathy for those who look like us. We feel empathy with those of the same color, same situation. And that has made a huge difference for where the money in charity goes. We’re talking billions of dollars here. It makes a huge difference for how wars start. Wars are often triggered by the suffering of people, and our empathy toward people. It has a huge implication on the criminal justice system. There’s a disproportionate number of minorities in the criminal justice system. It’s not some sort of accident.

Here’s a proposal I would make, and this at least expresses my feelings toward empathy. I would like the world to be such that people who use empathy in a political campaign or a social campaign are viewed with the same disgust as those who use racism or xenophobia.

TS: Wow.

PB: If you want a law to pass and you say “we need this law because there’s this suffering little girl and her life has been destroyed by government regulation or by Obamacare – or by the Republicans” – well, you can always find these little girls. People should say that’s bullshit, it means nothing.

TS: Okay, but what if you say “here’s a suffering little girl who got hurt by Obamacare, but there are fifteen suffering people that were helped by Obamacare. It seems like empathy is valuable as long as you’re not trying to exploit it. Isn’t what you’re describing the abuse of empathy rather than empathy itself? As long as people are feeling empathetic for everyone who is going to be affected by the policy rather than just one person who is going to be affected by the policy.

PB: Yes. But empathy doesn’t work that way, it’s not in our nature. I got into a discussion with a philosopher who asked what if you had super empathy that allowed you to feel the feelings of everyone affected by a policy. So, a million people here, a million people there, and also future people. And then use that empathy to guide you in your decision. If that were true, then I’d be fine with it. But that’s not how people work.

You know, your example is interesting. A little girl whose life is made worse by
Obamacare versus fifteen people whose lives were made better. I bet if that experiment was done you’d side with the little girl because empathy is ridiculously innumerate. The way empathy works is that one is more than two. There are plenty of studies that show this. So that’s the problem. Sure, sometimes empathy gets it right. Racism sometimes gets it right; flipping a coin sometimes get it right. But I would say, a pox on it!

2. My name is Inigo Montoya. You killed my father. Prepare to hire a good lawyer.

TS: Ok enough about empathy, let’s talk about revenge. You say in your book that our willingness to engage in third party punishment, even at cost to ourselves, is grounded in our instinct for personal revenge. This leads to the emergence of an impartial third party justice system, which you say is a good thing. And once that happens, then the revenge instinct is best transferred to this justice system.

PB: That’s right.

TS: You give the example where your car window was smashed. And you had a feeling of outrage but it was better to let the police handle that than for you to go around searching for the guy that did it. You also write: “If Inigo Montoya were around right now, he wouldn’t need to storm the castle to bring his father’s murderer to justice, the police would do it for him and fewer people would have to die.” But would that really be a good thing? For Inigo Montoya not to feel that instinct of personal revenge?

PB: You’d have fewer people dying.

TS: Right, but there’s something we lose too when we let an impartial third party justice system handle all of our business.

PB: Manliness?

TS: Not exactly, and I hate that word – it sounds like the opposite of what it’s supposed to mean. But we do lose some virtuous qualities, right? Inigo Montoya was very courageous, loyal to his father and to his honor. None of that can express itself if he just lets an impartial justice system punish the six-fingered man. The example I use in this context is True Grit. Mattie Ross is awesome. Part of what makes her awesome is that she personally sees to the avenging of her father’s death. She’d be a lot less awesome if she had hired some mercenaries to go get her father’s killer, while she kicked back in a rocking chair and waited for them to bring back to Arkansas. By going after the killer herself, she shows how much she loved her father. And that’s really the only way she shows it. She’s not an emotional or sentimental girl, but she shows the depth of her love through her willingness to risk her life to get the murderer. She
displays courage, loyalty, all those things that you can’t display if you just call the police.

**PB:** I think you’re right. At the end of Pinker’s book *The Better Angels of Our Nature* he talks about the decline of violence. One of the forces of the decline in violence was what he describes as feminization. Which is a decline in manly virtues that you’re talking about. And Pinker argues that, basically, we’ve shifted from a culture of honor to a “culture of dignity.” And that was one factor leading to the decline of violence. And then he addresses the view advanced by his colleague Harvey Mansfield, which is that much is lost when we lose the manly virtues. And this is the point you’re making and I think you’re right.

But I agree with Pinker’s response, which is that more is gained. So, I agree here. In a world where somebody murders your father, you track them down, bring justice yourself, you get to exhibit the virtues of courage, and loyalty and love that maybe you couldn’t otherwise by calling the cops. And I think that’s what fools us sometimes. Because even when the cops are available if neighbor molests our kid, we might go over there and blow his brains out.

**TS:** You think that’s wrong? To go and confront the guy that molested your kid?

**PB:** I think that, as a matter of broad public policy it’s wrong.

**TS:** Ok but what about as a matter of personal morality?

**PB:** Some philosopher, I forget who, said the best response to a personal insult is to give the person a sound punch in the nose.

**TS:** I think that was Derk Pereboom.

**PB:** Maybe. So let’s take that as an example. It feels a lot better, punching someone in the nose, rather than saying “Well we must agree to disagree” and walk away. Or to sue the person in civil court. But, from a standpoint of how I want the world to work, I would rather give up certain things and increase the chances of me and those I love living long, happy, unmurdered lives.

**TS:** Up to what point? I mean, you have two sons right? Say one your kids is getting bullied. Not in a way that’s life threatening or anything, just getting bullied. Would you want him to be a snitch or a tattletale and just run to the principal? Or would you want him to stand up for himself and try to fight the kid? Again, you’re in a nice Connecticut neighborhood. Nobody's going to come to your house and do a drive-by if your son beats the crap out of the bully. The worst that happens is he gets his butt kicked. But at least he stood up for himself. What would you rather have your kids do in those situations?
**PB:** I think the answer depends on what society we are in. It's more of a local question. In this school I give this answer, in that school I give that answer. So, if it's a school where the kids in essence are in a culture of honor – and I think many schools are – then I would advise my son to fight back because otherwise he would lose respect, standing, face. But if it's a school like my kids go to where there's a strong culture of dignity, a strong anti-bullying and anti-violence culture, I would actually want him to go to the authorities and get it resolved. Because it wouldn't ruin my kid's standing in that society. Now it's true, my kid then would not get to exercise his manly virtues. But really - what's he gonna do anyway?

**TS:** What's the downside of encouraging him to ignore the norm that encourages going to the authorities, and have his own code: “You messed with me, I'm going to stand up for myself. I'm not going to snitch unless I have to.” I mean what's the downside of that aside from maybe him getting suspended for a few days?

**PB:** So, there are certainly downsides - losing a fight could involve injury, and there’s some degree of risk, and being humiliated. Being suspended is not pleasant, so there are downsides. Now what are the upsides? Well, in some societies, the upsides are if you don't do it you can't hold your head up among these people and that's a huge thing. I agree with that.

**TS:** But even in this culture, there's a virtue upside maybe. A personal code that says: 'I don't care what the surrounding norms are: I stand up for myself.' Not for stupid things like someone cutting you off on the highway. And not when you risk getting killed or seriously injured. But in general, yes, I handle my own business.

**PB:** Which world would you rather live in? One where people settle disputes with their fists or one where people are looking to the culture, talking it out, seeking mediators, and so on?

**TS:** I don't want to live in either one of those worlds entirely. I want a world where there are limits. You're not going to risk your life to stand up to the bully, but you will risk getting your butt kicked. By the way, this reminds me of your tattling discussion in the book. You say that there are two reasons kids like to tattle. First, they show the adult how moral they are. And second, they satisfy your own sense of justice. But there's also a huge stigma for kids about being a tattle-tale, right? It's a pejorative word – it's shameful to be a tattle-tale.

**PB:** You know, you're right and actually it's an interesting question: why is that? One answer might be that we live in a culture of honor where you're encouraged to fight your own battles. But I think the stigma of tattle-tailing goes beyond that. Because we have third party punishment and it's still stigmatized. I think people see tattling as unfair.
**TS:** Unfair in what sense?

**PB:** I see you talk disrespectfully towards undergraduates, and so I go to your department chair. You're thinking that I wronged you, right?

**TS:** I'd want you to talk to me first.

**PB:** Exactly. That's the first thing you'd say to me, you should have talked to me first.

I think you raise a good point - I will concede that sometimes, personal revenge, personal payback, standing up for yourself in those physical ways, is actually a good thing for a person to do. It is good for your character, it exercises a mental trait that is well worth exercising. And, it's even good when other people witness it, and for the people who the bully tries to intimidate. I will concede all of that and I think some of these intuitions are at play when we agree that tattling is wrong, even in an academic setting.

But I still think on the whole that this culture of honor is something we have to lose, or we're better off losing at least in the physical way. I think it's similar to when people talk about rituals involving male/female differences, and chastity, and respect and all kinds of manners and so on. Some of these things, when viewed in isolation, are beautiful in and of themselves. And they're defensible in that sense, but the cost might be just too high. You can't have a society which has a strong ethos of nonviolence, which I really want, but at the same time allow kids to punch each other in the nose in school fights.

**TS:** I guess. But even in a non-honor culture, you don't want to encourage a society of tattle-tales either.

**PB:** You want to return to the time of duels?

**TS:** (Laughing) No.

**PB:** Anthony Appiah in his book *The Honor Code* argues that duels stopped because they started to look ridiculous and people make fun of people doing them.

**TS:** Right – now when we think of dueling, we think of French guys slapping each other with white gloves.

3. **Morality, Math, and Porn**

**TS:** You make two analogies which are really interesting. The first is the analogy of morality and mathematics, and specifically the zero concept. And then of course there's the one between morality and pornography. Can you explain both of those analogies?

**PB:** Do I make an analogy of morality and pornography? I don't remember that!
**TS:** Well, the analogy is in response to the idea that morality couldn’t have evolved by natural selection.

**PB:** Ok, right. The morality and pornography example is a response to an extremely bad, but extremely popular, argument which goes like this. We are often nice in ways that don’t benefit our genes in any imaginable sense. Ergo, morality could not be the product of evolution. And so morality must be a gift from God. This argument ignores the fact that there is a difference between the forces that cause something to evolve and the things that trigger it in its current day. An obvious analogy is pornography. So, pornography causes men to devote time, energy, sperm, and so on yet it has no reproductive benefits. Are we then forced to conclude that pornography is a divine miracle?

**TS:** Well, people say that this is a golden age for pornography.

**PB:** (laughs) Yes, I’ve heard that too. Anyhow, the example of zero is, as you point out, more interesting. It’s in a discussion of our moral responses to friends and family and strangers. The point of the analogy is that we have these natural moral responses to friends and family, which can be readily explained in evolutionary terms. But we are also nice to strangers. We give money to them, we volunteer to help them, we value their lives. Some of us even care about non-human animals, so much so we refuse to eat meat.

How do we explain this? I think morality towards strangers is like zero. The evolutionary story of how we come to make sense of zero is that there isn’t one. We’re not wired for it at all. Zero is a discovery smart people have made. It’s understood by children with great difficulty. And I think issues involving impartiality, fairness, kindness to strangers are the moral equivalent to zero.

**TS:** You say that the wrongness of slavery is a discovery in the same way that I guess zero is a discovery, or Newton’s laws is a discovery.

**PB:** To a philosopher like you I guess that sounds like a form of moral realism.

**TS:** Well, maybe. I would need to know the sense in which it’s a discovery. The case of zero seems like a discovery of reason. But with slavery, maybe it’s the Humeanin me, but that seems like it could be more of an emotional discovery, maybe one involving your arch-nemesis: empathy.

**PB:** No, I think it is a discovery of reason. I think we’ve come to these conclusions that say slavery is wrong because we have some sort of antecedent notions of what’s moral we can build upon. And another way this is a little like math, to some extent, is that it can bring us to interesting conclusions. And that’s what moral philosophers and ethicists tend to do. “If you believe this and this, well it follows that you should also
believe that.” People say, “You can’t argue about morality.” But of course we do that all the time. We point out inconsistencies—“you can’t say both that X is wrong if but then not complain when your favorite politician does it.” We work these things out. That’s what moral argument is.

**TS:** True, but there are often fundamental disagreements, right? You think it’s OK to favor your son’s welfare over a strange child in a foreign land. Peter Singer would say that’s irrational. Your son doesn’t feel pain or happiness any more or less than a child in Africa. Then you say ‘I favor him because he’s my son.’ And he presses you: “Why does that matter?” “Because he’s my son. It’s a basic intuition.” And on it goes. You don’t have that with mathematics, or at least you don’t have it to anywhere near the same degree. There aren’t mathematical disagreements at this fundamental level.

**PB:** I’m sympathetic to what you’re saying. Yes, there are some moral principles that can’t be justified. But there are all sorts of moral things you can make sense of. Why would you hold the door open for somebody? Well, duh, because that’s nice and helps them get through. Wouldn’t you want someone to do that for you? These are the sort of arguments one would be encouraged by. We never talk about them because they’re sort of obvious but they’re important and they’re significant.

Then there are the really hard cases, and I think your first example is the hardest case of all. And it’s one that I’ve personally struggled with. On the one extreme, I’d give everything to my family and nothing to anyone else. That seems wrong. The other extreme is that I’d give perfectly equal, and that seems wrong too. I feel we should live in the middle. I have no idea how to figure out what to do. I have gut feelings but I don’t know how reliable they are. Here’s an example. If I’m rushing to bring my son a new video game as a present and I rush by and somebody is drowning, I would stop and pull the child out of the lake even at the expense of letting my son wait. That seems obvious, right?

**TS:** Yes.

**PB:** Right. But if I was rushing to give my son a new heart, then I might run by the drowning person. And if I hear that you ran past a drowning person I’d say you did the right thing, I can’t blame you.

**TS:** Right and imagine if I dove in with the heart to rescue the stranger, because I was playing the odds. I figured I can definitely save this drowning person, but I’m not sure if the heart will save my daughter—you’d think I did something wrong, right?

**PB:** Yes. And then it’s complicated. What if it wasn’t the heart but it was the eyes? These are really hard questions. I think those are the hardest questions. But so much of morality isn’t like that. Take a case like animal welfare. People can argue about the treatment of animals, and the case is pretty strong that certain things we do cause the
animals great pain in order to give us unnecessary pleasure and that this is a really bad thing. And it’s true somebody could say, “I don’t give a shit about animals and I don’t think they matter morally at all.” And someone could reply: “They matter because they suffer – what makes humans so special?” And the argument can go on.

4. What is the Engine for Moral Change?

**TS:** Ok, but many cases of converging moral agreement aren’t the result of rational arguments. Let me take your own example. You say in your book and it’s a great line: “It might well be that the greatest force underlying moral change in the last 30 years in the United States was the situation comedy” You credit *The Cosby Show* for undermining racism, and *Will & Grace* for bringing people around to same sex marriage. I totally agree with you but that seems to be undercutting your point. The same arguments for treating gay people equally were available pre-*Will & Grace*. There haven’t been any new decisive arguments that convinced everyone. It was just people saying “Oh, okay, I don’t feel that this is disgusting or weird or creepy or unnatural anymore. They’re just normal people.” That’s not a discovery of reason, that’s familiarity and empathy.

**PB:** So here my view is interestingly nuanced! I think reason is a force that drives moral change.

But other factors surely play a role as well, both good and bad, and sometimes people who have arrived a conclusion often use these others factors to persuade people to join their side.

Exposure to others often has a good effect, and this is a good thing about television. On the other extreme, I think we entirely overrate, in our societies, the importance and value of honor, because it is so exhilarating to see it on TV. So, punching somebody in the nose for insulting you is a common TV way to resolve a dispute. And it’s always done by a hero and it’s always right. Well, you do it in New Haven and you’ll get a criminal record, you’ll be expelled from school, and people won’t like you.

So, sure there’s often a trickle-down morality. Some people come to moral insights and they get passed on to the rest of us. I think slavery is wrong. Do I think slavery is wrong because I worked it out myself? No. I consider it wrong because people say slavery is really wrong, and I picked it up from society.

**TS:** When you see *Roots* as a kid...

**PB:** That’s right. And I think gay marriage is right probably because everyone around me thinks gay marriage is right, so a lot of this is trickle-down. But this is true for science as well. This is why I think diseases are caused by germs, because people told me. I believe the earth is billions of years old because that’s what scientists say.
**TS:** Right but that doesn’t seem analogous to the gay marriage case. Why is the younger generation so much more accepting of gay lifestyles and marriage? It has nothing to do with people telling them homosexuality is okay. It has everything to do with them being exposed to and being around gay people and not having a problem with them. It’s just not an issue for my students at all – and I live in Texas!

**PB:** Let me push the analogy with science. There are two questions that have to be distinguished. The first is, why do so many people now believe in evolution? Not as many as it should be but a lot more than there was fifty years ago. That’s question one. Question two is, how do people discover that evolution is true and that species are evolving? It’s clear that these are two distinct questions. The former question might be because certain authorities say it’s true (and actually most people who believe in evolution have no idea what natural selection is). The second question of how we came to it is involves reason, experiment, and argument and so on.

I think it’s the same for morality. And so I think that if the question is: why do most people think gay marriage is okay, then the answer is because of the a-rational things we’re talking about. But how do we come to appreciate, as a society, that homosexuals have the same rights as heterosexuals? Here, there’s a lot of moral argumentation.

**TS:** You think that argumentation is a real or primary cause of the recognition of gay rights?

**PB:** Yes I do. I certainly think we’re capable of argumentation. You see it all the time, people do it tons. This is where social psychologists get it wrong because people argue about morality all the time, and we really exercise this capacity of reason. And moral change has a directional property to it. So you look at graphs of acceptances of gay marriage, or interracial marriage, and they just go up and up and up. And I think reason is the only process we know of that has that sort of directional property.

**TS:** What about familiarity? Couldn’t you make the same exact case for familiarity? Whenever there’s some sort of stigmatized group, increased familiarity will lead directionally towards a decrease in stigmatization.

**PB:** Yeah, I don’t want to be too quick to argue against that because the consensus probably is due in part to familiarity. You pointed out earlier that arguments in favor of same-sex marriage have been around for a very long time. This existence of these arguments is kind of my point. There are arguments in favor of it, there are reasonable people who deliberate and say that it might seem wrong maybe but it’s okay. And there are arguments back and forth. For slavery, there were people arguing that slavery is a good moral system and other people argue that it’s bad. And then over time the people who thought it’s bad won the day. Then their views spread to the culture at large, and this spread was probably aided by familiarity and other things.
TS: I wonder if there's an empirical question that the debate we're having turns on. Is it “what actually causes people to subscribe to a certain moral views?” Take animal welfare. Would our debate turn on the question of what really caused people to start caring about animal welfare? Was it because of the rational arguments from people like Bentham or Singer? Or was it the pictures in Singer's book Animal Welfare that made us say “holy shit, this is what they do in laboratories? That’s awful! This is what they do on factory farms? That's disgusting! We have to put a stop to it, I'm outraged!” If the latter is true, I'm right, and if the former is true, you're right?

PB: No, because I think they're both true. There will be a world fifty years from now where factory farming is considered morally repugnant. And we'll look back and say how could people not have known this? I don’t think the individuals in the future will have each made a reasoning move that we are now incapable of. They’re not smarter or whatever. Similarly there will be a world fifty years from now where everybody will agree about natural selection. So there are social changes that are not driven by reason. But, I think for the moral understanding to occur in the first place there has to be a catalyst or some kind of insight. There was a time when nobody would have thought slavery was wrong, and without the spark of some kind of moral insight or moral argument, we would go on having slavery until the end of time.

TS: So the argument sort of gets the ball rolling. And these other emotions can then be triggered....

PB: Triggered and exploited. I don’t want to be simple-minded about it, because there are cases where you could have moral arguments with bad conclusions. But science can end up giving bad arguments too. There are all these social forces and obstacles. But on the whole both moral reasoning and scientific reasoning leads to progress.

TS: But as you point out, reasoned arguments go both ways. You can reason your way towards conclusions, and the only way you know they're bad conclusions is because of how they feel. Would you agree that our core emotions provide an indispensable check on our reasoning processes? Jonathan Bennett describes a famous case of this with Hermann Goering who intentionally stifled his empathy because he thought it was his moral duty to exterminate the Jews. So Bennett’s idea is that you have your morality, your set of moral principles, but your emotions are the ultimate check on these principles.

PB: I strongly disagree with that. There are certain cases we can point to such as Goering, the Nazi doctors, where you have people saying “Wow this feels really repugnant, but I think my argument is sound” and they're just monsters. I agree with that. But from an evolutionary perspective these emotions have evolved from local circumstances. They are reliable to different degrees. I think disgust is not reliable at all
and empathy is in between, anger in between. But emotions are always limited, while reason is a ladder you can bring anywhere.

There were people in the 1980's saying "Well, you know, this gay rights issue is all well and good. But if you take their argument to the ultimate conclusion, then you'll have gay people marrying each other. And we know that's wrong, so hold on reason! Let's not be ridiculous." I think that the product of moral progress is that the gut feeling sort of fades, and we get an edifice up that pushes it away.

Now, where's this going to take us? Someone nasty could say to me, what do you think of bestiality? Consensual bestiality doesn't harm animals, but I think that's very gross. Do I think it's wrong? Reason, my utilitarian principles, could speak up to the point where I say certain forms of bestiality aren't wrong. And here I have my gut feeling say "Woah! I better get some better arguments." In the end, though, reason should always win. We should distrust certain gut feelings, and I think our rejection of gay marriage, and before that, interracial marriage, provides perfect examples of gut feelings gone wrong.

TS: But it all has to start with gut feelings at some point, right? Your gut feeling is what tells you that it's bad when people feel pain and good when people feel pleasure.

PB: Yes, that's right.

TS: So it seems like you want to be selective about which gut feelings you accept and which you don't. You might say that reason tells us which of them to accept.

PB: I think for any good moral psychology, there are two types of gut feelings. There are the gut feelings that are the grounds, the foundations of morality. And the truth is there's not much we can do about that. If you don't agree with me that causing unnecessary suffering is wrong we just can't have a moral discussion. A foundational intuition like that has to be taken as a given.

But there are other sorts of gut feelings, like gay marriage is wrong or interracial marriage is wrong. Because they are not primary, you can override them and work around them. This sounds messy, but as you know this is what moral philosophers do all the time. In the end, any decent moral philosophy will have to say "Well, this feels right but actually it's wrong." This process will push us in new directions, but it can't push us too far. If it ended up saying "killing and eating everyone you see is the ultimate moral good" you say "that's not moral, that's nuts."

TS: I agree with you on that. But it does seem somewhat arbitrary what we assign as the primary moral feelings and what we assign as expendable. As Jon Haidt loves to point out, for liberals the primary ones are usually the ones that are involved with harm or fairness and equality. Liberals assign the feelings about those categories as primary
and consider the other ones to be byproducts, silly, or at best derivative of the harm/fairness feelings. I just don’t know how principled that is. Because people who aren’t Western liberals regard feelings about family loyalty, group loyalty, respect within a hierarchy, and so forth to be just as deep as our feelings about avoiding pain or treating people equally.

PB: I think that’s right as a description of moral rhetoric. My strongly held moral views are foundational and undeniable, yours are sort of things you need to get past. Because mine are self-evident, and yours are just dubious! And I think you’re right about the cultural bias. As a liberal, I say we want to be fair to gay couples, who could deny that? And then a conservative might respond, well it violates certain traditions, and I say that’s ridiculous. That’s just an arbitrary feeling that you can shed. I guess in the end it might turn out to be an empirical question as to which intuitions fall into the different categories.

TS: So how can we make the distinction between what you call “foundational intuitions” and expendable intuitions?

PB: I guess I’d say that making the distinction between the two sorts of intuitions can often be done with common sense. In the end it’s an empirical question, but it’s an empirical question where sometimes it’s obvious what the answer is. So if you tell me you just feel in your heart that school vouchers are immoral, I’m going to say “Look I understand you have that intuition, but that can’t be a primary intuition. Rather, it has to be derived from other things.” Maybe you trust people who say so, or maybe it connects with other feelings - but certainly you wouldn’t say “Well the argument has to end there.”

But for standard reasons that philosophers point out, there has to be some intuitions that don’t derive from other intuitions. And I actually think it’s an interesting question whether or not you can talk people out of those. Maybe you can, particularly if they’re conflicting. Then the analogy would be to visual illusions. So, you see one line as longer but you’re persuaded that your perception is mistaken.

TS: There’s a disanalogy there though, which is that with visual illusions you can show me that I’m wrong. You give me a ruler, I do the measurements, and I say, “Oh you’re right, that’s an illusion.” But there’s no analogue for the ruler in ethics, is there? If there are two different intuitions that conflict, two intuitions that people think are foundational, it’s not clear at that point which one is the illusion.

PB: I guess it’s true. For intuitions of physical space and other things you could appeal to an external reality that you access through science and other forms of observation. You can’t do that for morality. So, if you say your moral intuition is that torture is always wrong and I say “Well what if torture saves lives in the long-run?” You could say well
that's data but that doesn't convince me.

I'm interested in hearing what you have to say about this. What really happens is: we have multiple moral intuitions and then we do this whole reflective equilibrium thing where we coordinate them together and try to resolve inconsistencies. And that's where the argument goes. Often it ends up with one person trying to convince another that their intuition is improperly derived.

**TS:** But what do you mean by 'improperly derived'?

**PB:** Hypothetically let's say I have an intuition that affirmative action is wrong because it is a form of discrimination, and discrimination is wrong. I can imagine someone arguing with me saying, "Well look you don't really believe that discrimination is always wrong." And then giving argument after argument until I concede, "Oh, okay I see some sorts of discrimination are okay."

And that's how moral arguments go. In other cases, moral arguments are plugged into factual considerations. You say "I think abortions are fine because a fetus feels no pain," but then look at this, I'm showing you a fetus does feel pain, and this might make you change your mind. A lot of moral arguments are sensitive to issues of consistency and a lot of them are calibrated to facts.

I'm really interested in the case of disgust. Something grosses me out, and I feel it's wrong. And I don't think you can ever really shake that feeling. But I think people could be persuaded that, on balance, disgust is the wrong way to make moral decisions and so is akin to an illusion. The sort of argument you'd make is to say, "Well, people had very strong intuitions, based on disgust, about interracial marriage. But surely you believe interracial marriage isn't wrong." And then this calls into question this mode of moral deliberation. Does that make sense?

**TS:** Yeah it does. You can even point this out for intuitions that seem primary. You can point out that it's inconsistent with another intuition they have that may be just as primary. And you can at the very least force them to reconcile it. You might even point to one that they themselves would admit was deeper, an intuition they were more strongly committed to without being aware of it necessarily. Is that the idea?

**PB:** These are questions I am struggling with. What does seem clear is that we do make moral arguments, we do persuade people morally. And this doesn't seem different in kind from other sorts of arguments people have that sometimes get productively resolved. Of course, there's a whole continuum of things. You and I could disagree about whether it's raining outside and we could go outside and look, that's on one extreme. That's easy. Another extreme is that you say *Straw Dogs* is the best movie ever and I said "meh" and maybe those are harder to resolve.
**TS:** That might be more in the middle. I was thinking that basic taste judgments are at the other end of the extreme. Like whether eggplant is delicious or horrendous. I think it's horrendous, I hate eggplant. My wife loves it - we're not going to resolve that one.

**PB:** But that's only a disagreement if you were to argue that eggplant is objectively horrible.

**TS:** I think eggplant might be objectively horrible but I see your point. You're saying there's a continuum for how objectively we tend to think of moral matters.

**PB:** Yes. And someone like Sam Harris denies the continuum. I think I'm reading him fairly to say that he believes moral disagreements are just factual disagreements in the end, and that moral views are wrong in the same way that statements of fact are wrong.

**TS:** But aren't you also committed to that view to some level?

**PB:** Yeah, to some level. If I think there's moral progress, then I think there has to be moral mistakes and moral truths that we can find through inquiry. But I also think there are points at which we actually can't - which are very hard to resolve that way. Appiah had a very good review of Sam Harris' book, where Appiah says "Look, here are some examples of some things that can't be resolved through scientific argument." And he gives versions of the repugnant conclusion - is it better to raise everybody's happiness a tiny bit while one person's happiness drops a lot? And he points out correctly that there's no factual way to answer this question.

Actually, the examples which strike me the most, as we talked about earlier, are the balance of our obligations toward family and to strangers. So Harris says, I think correctly, that we care way too much for those close to us and not enough for those in faraway lands. Okay, I agree with that. But if you push him, he can't tell us with a scientific analysis exactly how much we should care about people in faraway lands.

**TS:** He can't even come close, right? That's a great example of something science can barely shed light on.

**PB:** Right. You can imagine that we have all the data on exactly how happy people will be if you do such and such, how happy my son will be if I send him to a good school, how happy Africans will be if I save their children's lives. Every molecule is known, every data point is exact. Still, I don't know what to do.

**TS:** Exactly. OK, last question on this, because I'm still trying to figure out ultimately where you stand metaethically, in the debate over whether morality can be objective. Would you call yourself a pluralist? You seem willing to allow for a good amount of flexibility as to what can be called morally good. And everything you've said is consistent with the idea that there are multiple moral truths or moral systems that are
jointly incompatible. A single individual could coherently hold any one of them but not all of them.

**PB:** That's more of a psychological claim than a metaethical claim.

**TS:** Well, it's metaethical in this sense. If I'm right, there would be no way of (a) persuading a person who embraced a different morality and also (b) providing a principled justification for one over the other. There would be no way to adjudicate between them. You see what I'm saying? It's a psychological question whether they could be convinced one way or the other, but it's a philosophical question whether one is more justifiable than the other.

**PB:** I honestly don't know. I'm less of a pluralist than someone like Jon Haidt. Jon has all these different moral foundations, and I think his empirical psychological claim is largely true. And then he's a pluralist about it—he thinks that these are different ways to live life well. My own view is that you can look at a moral foundation, a psychologically valid moral foundation, you could track its evolutionary history, developmental history, and then you can say "This really sucks." I think people are often tempted by a version of the naturalistic fallacy, which is that if something evolved through natural selection and cultural selection, then it's somehow good. I again think there are certain moral sentiments which you can look at in a cold-blooded way and say yeah it served a function but now it's actually not good, we'd be better off without it. And, here's the controversial point, I think that upon reflection any rational person could come to agree with that.

So, consider female genital mutilation - there's a cultural history, to some extent it might even be a cultural adaptation and thus serves a purpose for having it. But now we can look at it and say that actually it does not make the world better, it makes the world worse. Same for slavery or whatever. More generally, I think the whole foundation of disgust and sanctity is like that.

**TS:** So you would describe yourself as less of a pluralist than Haidt in that you think there's more work you can do beyond the foundations. But that doesn't mean you're going to arrive at a single truth or single correct moral worldview, right?

**PB:** That's a really good question, and I just don't know. I'm not just saying I don't know which one is right—consequentialism or all the forms of deontology or virtue ethics or whatever. What I don't know is whether or not there is any way to make this sort of choice.

**TS:** What does that depend on, empirically, for you?

**PB:** My inability to answer that question is why I don't know. So, it goes back to what we were saying about Sam Harris. Not only don't I know how much of my money I should
give to charity, I don’t even know how I could know.

**TS:** We can’t even conceive of an experiment or any evidence that could answer that question.

**PB:** Right. Because Harris thinks all you have to do is accumulate data on other people’s happiness and your own happiness. But that can’t be right. Because even with the full data, you can draw up these different curves of how much happiness I have, my son has, Africans have, and so on, but I still won’t know what to do. But it’s also because while I think I owe something to strangers, the idea that I don’t owe more to my children seems so unnatural that it doesn’t seem to be morality at all.

**TS:** And yet you think an impartiality principle is one of the foundations of morality, right?

**PB:** Yes I do. But the impartiality principle, I think, is not inconsistent with saying we have different obligations to those close to us and those far away.

**TS:** It doesn’t commit you to favoring everyone’s welfare exactly identically?

**PB:** Right.

**TS:** So the empirical question all this seems to rest on is this: with full information, would a rational person, even one from a different culture with different practices, agree with you?

**PB:** Yes. And put as a claim about actual people it’s probably wrong. I’m a Jew and so if you tell me that male circumcision is stupid, I might say well screw you, you anti-Semite! But that just goes to show that we’re not perfectly rational. So I guess maybe my claim is that, if we were all perfectly rational, perfectly smart, unbiased, moral agreement would shoot up. But I also can see that it would not perfectly converge. You might think that you should give 40% of your money to charity, I think 20%. And we might just disagree and not know how to resolve that.