# Philosophical Profiles

# David Shoemaker

Associate Professor in the Department of Philosophy at Tulane University and Faculty Member of the Murphy Institute

**IN BRIEF**

David Shoemaker is Associate Professor in the Department of Philosophy and the Murphy Institute of Tulane University. His areas of interest lie predominantly in the topics of personal identity and moral responsibility. He runs the biennial New Orleans Workshop on Agency and Responsibility, out of which come the Oxford University Press series Studies in Agency and Responsibility, for which he is the editor. He has written a bewildering profusion of articles for leading journals, and three books, the most recent of which is *Responsibility from the Margins*, hot off the presses from Oxford. No library is complete without a copy. David will fight anybody who questions the notion that Peter Strawson's article "Freedom and Resentment" is the greatest article in the last 60 or so years. Don't get him started on psychopaths.

**DETAILS**

Simon Cushing conducted the following interview with David Shoemaker on 27 July 2015.

**CITATION**

# David Shoemaker

## a philosophical profile

**SC: How did you discover philosophy, what made you think 'ah philosophy is what I want to spend my life doing'?**

DS: As I'm sure it's been for many people, it's a product of not wanting to do something else, so I went to college specifically to study pre-law, and I was going to be a lawyer. And I was told that majoring in philosophy would be a good idea for that, so I double majored in philosophy and history. But the very first philosophy course I took was a slap across the face. I realized that this was really a challenge to me and it was a challenge that I loved. But once I graduated from college I still was intending to go to law school, and then I worked at a law firm as a word processor for two years and when I saw that they had bathrooms and showers there, that they intended for the lawyers to use because they were supposed to be sleeping there and working seventeen to eighteen hour days and everybody I talked to was basically pushing commas around and seemed miserable, I said this is not the life for me. I looked around and decided to go to graduate school in something that I really did love, which was philosophy.

**So what was the first issue that really engaged you?**

I wanted to do applied ethics; I wanted to explore topics like abortion and legality of pornography. So I was interested in a number of applied issues but I quickly realized once I got to graduate school there were a lot of theoretical issues that I knew nothing about that seemed to ground the work that was being done in applied ethics so I got more abstract and finally starting to work my way out of that back into some more applied realm, but that's what I wanted to do. To do that I realized there was a lot of background theoretical work that had to be done if I wanted to do a good job at it.

**I think that's a sign of someone good at this stuff, is when they realize how much there is to do. It's whenever a student thinks this is pretty easy and there is very little to be said when you realize ok, it's passing you by…but when you are confronted with the enormity of things…**

So there are two reactions one might have. One is you realize how much more work there is to do and you put it into context and the other is to quit, so there's that temptation often.

**What was your dissertation about?**

My dissertation was on personal identity. I had become very intrigued by that. I had a great professor by the name of Greg Kavka.

**Everybody I know who knew him seems to have worshipped the guy.**

# David Shoemaker

He was just a wonderful man, died far too soon. It was in his first year of my dissertation as my advisor that he died. That's when Gary Watson stepped in. He had revealed to me the stuff that Derek Parfit had been doing in Reasons and Persons, and I fell in love with it. It was also interested in applied ethics because what Derek Parfit does in section three in that magisterial Reasons and Persons is to explore issues of abortion, responsibility, desert, compensation, these are applied issues but through the lens of a more theoretical lens of personal identity. That's what I wanted to do and so the dissertation was on personal identity as it pertains to various, of our more specific practical concerns.

**So you started out writing about personal identity, but it pretty quickly bled over into ethics as in your book Personal Identity and Ethics. But before you get into the ethical stuff, how would you summarize the state of the field of the philosophy of personal identity? Because when I went to graduate school it was just after Parfit and for a while there it looked like Parfit had capped it off—there were all these problems with the Lockean view and Parfit came along with this totally radical response, in some sense replying to the Lockean view and in another sense saying "Why should we care about personal identity anyway?" But it seems recently, in the past decade and a half, there's a whole new burgeoning of the field: would you say that's right?**

I think that's right to an extent. What we've got is some very smart people responding with a biological criterion of identity, which is intended to, I wouldn't say refute the challenge of Parfit, but basically deflect it. To say that what matters for people is the psychological relation to something, so something pertaining to a psychological criterion of identity and that captures a lot of our practical concerns that are person-related. But when it comes to the metaphysics of our identity, then what we are is essentially biological creatures so what constitutes the preservation of our identity across time is just biological continuity and so there's something deeply attractive about that to me, it seems right. The reason I say I'm not sure how burgeoning it is, is that in a way there's a kind of stalemate that's still at work between the two big criteria of identity. You still have people who are holding onto a kind of psychological view, Harold Noonan and to some extent Lynne Rudder Baker and Jeff McMahan (his is an interesting kind of hybrid), and then Eric Olson and others, Paul Snowdon, holding to this biological criterion and there are some interesting exchanges that are going back and forth between them, but to my mind many of the issues have been settled to my satisfaction, that is to say I think there's something like a biological view of metaphysical identity that's correct because I think that all of the relevant ethical and practical concerns are utterly independent from questions of our metaphysical, numerical identity. The one exception to this is the very recent work by Marya Schechtman, which I think does advance the debate in a really interesting way. I'll be teaching her book this fall. What she's doing is exactly what I think is the next natural and maybe endpoint step, although I should never say that talking to other philosophers, but it's to say that there's something very right about our identity as organisms but it doesn't capture the kind of special object that we are that's distinct from animals. And we are human beings, so it's a kind of anthropological view of our identity. Some of us become persons, sure, but there are identity conditions that must attach to people who are not full-fledged persons with the full-fledged psychological capacities that we've got that may not achieve the capacities for responsibility and such, but nevertheless are the object of a whole range of our practical concerns. To the extent that we are looking to the object that perseveres across time, we need to look at

> 66 To my mind many of the issues have been settled to my satisfaction, that is to say I think there's something like a biological view of metaphysical identity that's correct because I think that all of the relevant ethical and practical concerns are utterly independent from questions of our metaphysical, numerical identity. The one exception to this is the very recent work by Marya Schechtman, which I think does advance the debate in a really interesting way. 99

human beings in their full anthropological glory. It's a beautiful theory, to my mind it's probably—that's the one, that's the one, I think.

**I have to say I understand from the point of view of metaphysics why Olson tries to make this the issue of ontology the central issue, but to me it seems to me it's a parallel issue and the issues we really care about are the things that we say what matters; the responsibility issues and things like that. I think that actually you can have both of these running parallel and they're talking across purposes. Olson's argument about the "too many thinkers"—who is it thinking, the person or the animal? I don't see the knockdown force of that. What do you think?**

I agree with you on that. I always found that to be taken by some to be a knockdown argument and I'm not sure that it is. It's a kind of appeal to oddness and I'm not sure how much force those kinds of things have. The idea is that you have a person and an animal sitting there, you're right there looking at a person and an animal, they're both obviously thinking and isn't that odd? There's a puzzle there but I'm not sure what force I'm supposed to take from that. I think the arguments in favor of a biological criterion or a more human-oriented criterion are supposed to be cumulative. We've got that kind of argument, but you have to couple that with the appeal to the argument about surely I was a fetus once and I would be someone in a persistent vegetative state, that's certainly the way we talk, committed to something like that, and we have to take that seriously as well. There are ways of responding to both of these arguments on behalf of the psychological continuity theorist, but there's also a kind of force to our being continuous with the rest of the animal kingdom and the kind of identity conditions that apply to non-human animals, you would think, would apply to us as well. Of course there are going to be added wrinkles to that, but the added wrinkles are the higher consciousness, it gives rise to the practical concerns we've got. But that's where I want to come in and say those we can treat separately because they're not actually grounded on identity at all; they're grounded on a variety of other relations. So that's why I'm perfectly happy to grant the metaphysical identity theory to the biological theorist or to someone like Schechtman, a more human-oriented theory. Nevertheless I don't think that the relations, the metaphysical identity relation they're pointing to, capture the variety of our practical concerns.

**So if an undergraduate came up to you and said, "Will science let me live forever?" What would you say?**

No, no—you need to give up on that! But you might be able to get something that is, to borrow a Parfitian phrase, "just as good as" ordinary survival. You might get the persistence of a certain kinds of psychological relations that will enable various of your practical concerns like anticipation or self-concern of a sort that might be grounded if, let's say the singularity comes and we've all become one with the Borg.

**To paraphrase Woody Allen: I don't want to live on through my clones, I want to live on in my apartment.**

Yes, so there is that aspect of it but to the extent that there's some kind of sufficient persistence of, say, your first person perspective somehow, or some entity that seems

> **"** The idea is that you have a person and an animal sitting there, you're right there looking at a person and an animal, they're both obviously thinking and isn't that odd? There's a puzzle there but I'm not sure what force I'm supposed to take from that. I think the arguments in favor of a biological criterion or a more human-oriented criterion are supposed to be cumulative. We've got that kind of argument, but you have to couple that with the appeal to the argument about surely I was a fetus once and I would be someone in a persistent vegetative state, that's certainly the way we talk, committed to something like that, and we have to take that seriously as well. **"**

to remember all your thoughts and experiences and they were caused in the right way, that may be sufficient to ground some concerns that we have.

**They get my endowment; they get the money I buried in the backyard.**

They might, yes.

**All right: what is so amazing about P.F. Strawson's "Freedom and Resentment"?**

In the preface of my book I […] have often called it the greatest philosophical article of the twentieth century. And I have colleagues who make fun of me for this.

**You're a fanboy!**

Yes, I am; it's really what got me into the realm of philosophy that I have pursued for the last ten or so years. What I think is special about this article is that it was a complete game-changer. We throw that word around, but it was a genuine game changer for the way of philosophy of action and moral responsibility and freedom had been done. In that respect it was like Rawls's Theory of Justice. Utilitarianism had been the story of the day until Rawls had tried to resurrect a Rousseauian-Kantian view of political liberalism. For theorists at the time working in free will and moral responsibility, there were just two options, and it was a kind of compatibilism between freedom and responsibility and freedom and determinism that said that the two were compatible and so what we're doing when we're responding to people with these responsibility responses, like resentment and punishment and so forth, is really just a kind of social efficacy. Punish somebody and to that extent it's going to change the way they think in the future, it's going to deter other people from engaging in these kinds of things. Determinism can be true; their behavior could be completely determined, but that's okay. On the other hand you have people saying, well no, no, no in order for you to deserve punishment it's got to be the case that you are an agent that is outside the realm of natural event causation. These are the only two options and what Strawson did was present a genuinely compatibilist view of responsibility. That is, it doesn't matter whether or not determinism or indeterminism is true, it doesn't even matter what determinism is. Instead, if by paying attention to our natural human practices with one another we can see the sorts of things we are committed to in these responsibility practices and it—to question that from outside of the realm of the praxis by saying are they justified, if determinism is true or false, is just the wrong kind of question to ask. It was revolutionary in that it moved forward a kind of stalemated debate about the relation between freedom and determinism in a very valuable way and also something I think was very valuable was that it took very seriously a Humean approach to responsibility by paying attention to the real psychological details of the way in which we interact with one another. I think a lot of people found that very refreshing, very insightful, and so it's become certainly the most-cited article in this field by far.

**Which is kind of odd because when I think of Strawson I think of him as a philosopher of language. I think of his major article as "On referring" or something like that. And he didn't really make many forays, basically two, into these issues, and it also seems to me that it's an article that I'm not so sure was seen as so revolutionary at the time; it seems like it's grown in reputation—people have said, "Oh, look what he did back**

> 66 What I think is special about this article is that it was a complete game-changer. We throw that word around, but it was a genuine game changer for the way of philosophy of action and moral responsibility and freedom had been done. In that respect it was like Rawls's Theory of Justice. Utilitarianism had been the story of the day until Rawls had tried to resurrect a Rousseauian-Kantian view of political liberalism. It was revolutionary in that it moved forward a kind of stalemated debate about the relation between freedom and determinism in a very valuable way. 99

**then". It's like a hit that maybe made it to the top-ten but now it's twenty years later and suddenly it's number one.**

It was not cited very much for twenty or so years after its publication. Jonathan Bennett wrote a very influential article about it around 1980, I think, called "Accountability" and a few others referred to it, Gary Watson did an interpretive essay on it, I think, in the late'80s, 1988 probably. After that it started gaining a lot more attention. I don't know what the precise spark was, maybe it was early social media or something, "You've got to see this thing!" And you're exactly right—there's a biography about Strawson that doesn't even mention his work in "Freedom and Resentment"; it's all metaphysics that he had done. He had only written two articles on ethics and responsibility and he said "This is all I have to say on it, I don't think it as hard or as interesting as the work that I focus my attentions on in other realms," and Lucy Allais has actually said that he told her that he wrote both of those articles in ethics, including "Freedom and Resentment," in one draft. Which may explain a lot.

**It kind of helps: if you want a paper that people will go back to again and again, don't be too clear. Allow for plenty of interpretation. And it does seem to me that part of the paper seems rather dated and of its time, like the stuff about determinism, and that seems to be the stuff that people don't really talk about so much. Like the soft determinist versus the hard determinist—people don't really care about the "freedom" part; they care about the "resentment" part. What seems to be influential is his talk about reactive attitudes and his saying "It's a pity that people don't talk about the moral sentiments anymore." And people said, "Well we shall!"**

I think the respect in which the features that you're pointing to might be dated for some people is a direct function of that article. I think there were lots of people who were convinced by both that and Harry Frankfurt's early work on alternate possibilities that, okay, maybe the freedom or the metaphysics of freedom matters less here than the practices of moral responsibility and the attitudes that we have there. That being so, there are plenty of people who are still interested in the metaphysics of freedom and are doing interesting work and are interested in trying to reconstruct the Strawsonian argument in a way that works, because lots of people think that it doesn't work. But, as I say, for many that doesn't matter. We turned and we focused on the moral sentiments, the reactive attitudes, that's where the real action's at.

**It certainly gave people the license to talk about responsibility without discussing the metaphysics of freedom.**

Exactly.

**It's funny, there's such a disconnect between various sub-spheres of philosophy, for example, if you do philosophy of religion at all, it's taken as a given that you've got to have libertarian free will or you can't have morality or love or anything like that. I don't know if that's because of the issue of the problem of evil or whatever, but that's just taken as given—you must have libertarian free will. So you cannot have responsibility or any kind of moral value without that, whereas if you're doing ethics, we advertised for a post here in ethics and that's how I know how popular**

**Strawson's article is because everybody was talking about Strawson. It's like, wow, this didn't happen when I was in graduate school.**

It wasn't the case when I was in graduate school either. I really discovered it and its value several years after. But you're right about the necessity of libertarianism in certain sub-disciplines. And part of it is that there are many philosophers of religion who are going to be concerned about what Galen Strawson called "ultimate responsibility" which is going to lean heavily on your desert, the desert of hell, for example, that's an eternity of punishment so there better be some serious desert there. For that to be the case, that's heavy-duty desert, what Derk Pereboom calls "desert in the basic sense" that just for an action or for a series of actions, it would be appropriate to throw you in hell for all of eternity. Then you've got to have very robust kind of freedom and it's got to be contra-causal. I think that's what's driving that.

**Your more recent stuff—let's segue—you start out in personal identity and now your new book, Responsibility from the Margins, is entirely about responsibility. What drew you from one to the other because there certainly is an overlap for you, so how'd that happen?**

It's in part the writing of the first book on personal identity and ethics and one chapter in there is on responsibility. I remember struggling at the time with how to say what I wanted to say about responsibility and I realized there was a range of issues that needed to be discussed. I started to gradually get into responsibility. My thesis advisor was Gary Watson and when you have Gary Watson as an advisor, it seems like he's cornered the market on the truth. That's how it seemed to me.

**Never write your dissertation on what your adviser does. Either you just agree with them or what can you do?**

That's right, or minor little janitorial work. I wasn't interested in that. But gradually I started to see there was a little room and it was partially as a fuller discovery of the Strawson article. I had started to teach the Strawson article and also to teach issues in responsibility more generally. Then I became very interested in marginal agents. And by marginal agents I mean people who we think have a foot in and a foot out of the moral responsibility community, and I really found myself drawn in by the case of the psychopath. And so there was a literature that had begun around this time—this is the early to mid '00s—and I was fascinated by this literature. I won't say any more about this, but a psychopath entered my life in a certain way and so I became personally invested in this literature, and thinking my way through the issue of, is this kind of person, who has some clear incapacities, but who also clearly seems cruel and manipulative, responsible? And from there I started to think about, what about other kinds of marginal agents, those with intellectual disabilities, and those with autism? There was a paper I wrote in 2007 ("Moral Address, Moral Responsibility, and the Boundaries of the Moral Community," Ethics 118) where I tried to explore some of these issues for accountability, and as I thought more and more about it, it became clear to me that what was being forced upon me was a kind of pluralism about responsibility, that there are multiple types of responsibility and our responses to these marginal agents, I think, reveal that to us. It started with the thoughts about personal identity and responsibility, and as you say that's still there in part because one realm of responsibility has to do with attributing

> **"** I became very interested in marginal agents. And by marginal agents I mean people who we think have a foot in and a foot out of the moral responsibility community. As I thought more and more about it, it became clear to me that what was being forced upon me was a kind of pluralism about responsibility, that there are multiple types of responsibility and our responses to these marginal agents, I think, reveal that to us. **"**

certain character traits to us and character traits require a certain amount of persistence over time. The issues are absolutely related. I was also—this is going to sound bizarre—but I was intrigued by something that I wrote in the earlier book that I looked back on that said in the case of responsibility what we're doing is negotiating in certain cases, whether responsibility obtains. One kind of case is to think about a brother and a sister sitting at Thanksgiving twenty years after the brother has pushed the sister out of a tree when they were seven or eight years old and she still blames him for that and he says, "Oh come on—that wasn't me!" And he's not trying to escape responsibility, but he just wants to say, "I've moved on, I'm not that kind of person anymore," and I wondered how do we resolve this kind of issue? The thought is well maybe we're talking past one another in a way, or that she's appealing to one kind of responsibility, he's appealing to a different kind of responsibility. It was those kinds of thoughts that raised the issue; maybe we've got multiple kinds of responsibility at issue here.

**Now what's distinctive about a Strawsonian approach is it says: let's focus on what we do; let's look carefully at the feelings that we have towards other people. Of course a critic of that approach is going to say, we're stupid, we don't know what we're doing, we're wrong about so much. Why should something so important as responsibility—certainly if you're into the philosophy of religion and you're looking from a God's eye point of view of what people deserve—why should we care about these kinds of responses? For example you see people react totally inappropriately, it seems, to children, or to autistic individuals. They have the same feelings, they do resent them, but we want to say, yes you have those feelings and they're misplaced and the reason why they're misplaced is because the objects of the feelings objectively don't have the capacities that are required. So a critic of the approach would obviously say you shouldn't start with the feelings.**

Right that's absolutely the right kind of worry to raise here. There are many things that need to be said. One is that for a Strawsonian theory to be at all plausible you can't just be doing ethnography. I'm not just surveying the landscape of our responses and we say, and now we're going to build the contours of responsibility directly from those responses. There's got to be some normativity built in there. We can say of someone exactly as you put it: "Well no, you're just wrong to respond in that way to a child or somebody or who is severely intellectually disabled or autistic". How do we do that but hold onto the strategy of starting with the responses and the practices? One argument in favor of starting with the responses is that distinct from many other fields in philosophy, I think the study of moral responsibility is so inexorably intertwined with our responses that you've got to appeal to them on some level, at some fundamental methodological level. I've often poorly, lamely joked that you cannot spell "responsibility" without "response." But in this case it's really true that when we're theorizing we're trying to put together a theory of responsibility, the way we do so is by appealing to cases and leaning on our responses to those cases. How would we feel if somebody did this but nevertheless they were under severe stress, say? So I think you've got to appeal to the responses. Now the question is: do we take a Strawsonian approach where the responses are where we in fact start and we build up our theory from an appeal to the responses, or do we use the response as a kind of epistemic guide to the true facts about responsibility that are in fact response-independent. I try to stay neutral on that question in the book and just want to say that regardless of which way you go you've got to take the responses seriously. Now we don't just take the responses seriously.

> ❝ The case of responsibility what we're doing is negotiating in certain cases, whether responsibility obtains. One kind of case is to think about a brother and a sister sitting at Thanksgiving twenty years after the brother has pushed the sister out of a tree when they were seven or eight years old and she still blames him for that and he says, "Oh come on—that wasn't me!" And he's not trying to escape responsibility, but he just wants to say, "I've moved on, I'm not that kind of person anymore," and I wondered how do we resolve this kind of issue? The thought is well maybe we're talking past one another in a way, or that she's appealing to one kind of responsibility, he's appealing to a different kind of responsibility. ❞

It can't be that we just take a poll and find out how people in fact respond. Instead we've got to have some account of what makes the responses fitting, what makes them appropriate. What are the correctness or incorrectness conditions of those responses? To do that kind of work it's difficult and that's actually the project I'm working on now. I think what we need is simply that there are certain kinds of responses that we recognize generally to have certain kinds of constraints on them. And if we can't precisely identity what exactly those constraints—those normative constraints—are, we can get a rough picture of them. And that may be enough to get us to enable us to move forward and theorize adequately about responsibility.

**What would you say if someone said, "But wait a minute: by using that very term 'fitting' aren't you implying that the most important standard is the objective one, it's 'fitting' when they have the appropriate features. So it's the features that come first, not our response to them?"**

This is exactly the project I'm working on now and I'm very excited about it. I want to run an analogy in response—an analogy to humor. What is it that makes something funny? Think about the variety of things that counts as funny. We've got slipping on a banana peel, a good pun, two ten year old boys farting, two ten year old boys laughing about the fart that they've just produced. All of these things are hilarious; it can be somebody falling down, it can be a well-crafted joke, it can be people's bad reaction to a well-crafted joke—so I'm thinking of The Onion, who had some headlines about Harry Potter being Satanist, that was hilarious in and of itself, but what was also hilarious was that there were thousands of people who took that seriously and were forwarding it to their friends and telling them not to read the Harry Potter books—that is hilarious too. You've got this huge range of things that we find funny. What united them all under the rubric of 'the funny'? There are a variety of theories, and I'm just going to really gloss over, that all have counter-examples. One plausible response to the fact that they've all got these counter-examples is that what unites them under the rubric of 'the funny' is that they're all just things that we humans tend to laugh at, tend to respond to with amusement. Now not everything we respond to with amusement is funny and some of the things we don't respond to with amusement are in fact funny. But we've got ways of talking with one another about this, where we're appealing—people have bad sense of humor, people have good senses of humor. But this is a matter of some negotiation with one another trying to identify what it is that in fact is appropriate to respond to with humor. Here's the way that that structure would go: we've got what counts as funny is determined by our range of amused responses, suitably constrained in some way; you've got to have a good sense of humor, a sophisticated sense of humor, certain kinds of understandings of the world, and what that kind of sensibility responds to is what counts as 'the funny.' Now that has an objective status at that point so when you fail to laugh at what that thing is, I can say, "But it's got that thing!" and when I appeal to that thing it sounds response-independent, but what put that thing in the realm of the funny in the first place was that it was the sort of thing to which our sophisticated humor sensibility responds to with amusement. The exact same story is going to go with responsibility. Take something like anger, which we think is a blaming-response, when it's constrained to other agents. Think of the variety of things that we respond to with anger, it can be somebody stepping on your foot, somebody looking at you in a certain way, somebody treating you with condescension, somebody harming their child. There are a variety of things here, so what unites them all under the rubric under

> **"** Take something like anger, which we think is a blaming-response, when it's constrained to other agents. Think of the variety of things that we respond to with anger, it can be somebody stepping on your foot, somebody looking at you in a certain way, somebody treating you with condescension, somebody harming their child. There are a variety of things here, so what unites them all under the rubric under the anger-worthy, of the blame-worthy? There are various theories; there are reasons-responsiveness theories, deep-self theories, quality of will theories, and I think each of them has certain counter-examples in the same way that the various objective theories about humor do. I think the way to go is a response-dependent route; what unites them all is that these are the things to which we, having a suitably informed sensibility, respond to with anger. **"**

the anger-worthy, of the blame-worthy? There are various theories; there are reasons-responsiveness theories, deep-self theories, quality of will theories, and I think each of them has certain counter-examples in the same way that the various objective theories about humor do. I think the way to go is a response-dependent route; what unites them all is that these are the things to which we, having a suitably informed sensibility, respond to with anger.

**Maybe what you could say is what you're advocating is a reflective equilibrium. We say we're going to examine humor. So what's funny? The thing that explains why we as humans have a concept of funny in the first place is the fact that we laugh at things. So you can say that's got to be the starting point. Now it can turn out that we find out that there's nothing that they all have in common, but there are interesting core groups. And we can say that's the real funny, and funny isn't everything we laugh at, the really funny is the important core. So we can say that even though we were led to find this core by starting with the laughter, with our response. And we can bounce back and forth between that, and then we can say the starting point came with the laughter, started with the response. Is that what you would advocate?**

That's not what I would advocate. I think that's a more plausible view of what I advocate. But I'm going to stick to the more radical response-dependent view. I think the corrections that we engage in—it depends on how you fill in that reflective equilibrium view, maybe these are more compatible than I'm thinking that they are—the kinds of corrections that we do are not corrections toward a core, even though some people have thought this. So here's an example that Michael McKenna gives that is meant to cut against a response-dependent view of responsibility: we used to think that dyslexic children—we didn't know what dyslexia was—we thought that they were lazy, so they got lots of blame for not working hard enough in school, and so forth. And then we discovered they have this condition, dyslexia, and it makes things much harder, and so we changed; we changed our practices so that we know we ought not to blame dyslexic children. The thought behind that might be that, well what we're doing is engaging in some kind of reflective equilibrium, maybe we see that we respond to a certain kind of future people, but then we're correcting when we find out that these children don't exhibit that feature, something like that. What I would like to say instead is that what we've identified as a pure product of appeal to our responses is something like quality of will. And I think that fractures into multiple ways, but let's say in the case of dyslexic children that they didn't exhibit a certain kind of quality of will—that remains fixed. And then we discover certain facts about them and we realize, "Oh, okay, they weren't expressing a nasty quality of will, or a lazy quality of will," and so the attitude that we would have thought to be appropriate to a poor quality of will are no longer fitting. Now that quality of will, what the target of our responses is, hasn't changed at all; instead our responses change in light of certain empirical facts. I would want to say that not just the core but what counts as the range of the responsible, or the blameworthy, is entirely a function of our various attitudinal sensibilities, and any corrections that occur are in light of new empirical facts.

**It seems like you want to say, when you say, "our responses," one way in which you could say that the responses can be criticized is by focusing on a species-level**

**of response. So you could say: "Your response is wrong because it's not the typical human response".**

That seems right to me, and that's why I want to focus on a very distinctive subset of emotional responses that are responsibility-y. These are the universal human sentiments: I want them to be anthropocentric—it's relative at the species level.

**Like Hume.**

Yes, that's right.

**Maybe this would be a good time to have you lay out your tripartite theory.**

Most people, when they are exploring the nature of responsibility and blameworthiness, are appealing to, as Strawson did, a very tiny subset of our responses, and they are usually the triumvirate of resentment, indignation, and guilt, with occasionally gratitude. Mostly focused on the negative and on a very small subset. The first thing to do is to identify the much wider range of emotional responses that we have to people, people's expressions of practical agency—which is what I'm thinking very roughly as the rubric of responsibility. And so these include things like: of course, resentment, indignation, and guilt, but more generally anger, gratitude, esteem, disesteem, approval, disapproval, pride, regret, admiration, disdain, contempt, and the many, many others. Then what I attempt to do is identify from universal syndromes and it seems to me that there are three distinct categories (and these are pairs). The first pair is admiration and disdain, which are universal—and I'm thinking of pan-cultural here, universal sentiments, admiration and disdain and in particular a version of admiration and disdain, that requires me to say something first about the nature of emotions. Generally they have a triple syndrome. They include typically a kind of affect, an associated thought (not a judgment necessarily), and most importantly, a motivational impulse. And so I want to define these emotional universal sentiments in terms primarily of their motivational impulse. And so the motivational impulse that's definitive of admiration of a particular sort of other agents, namely, is a kind of motivation to self-improvement or emulation. Disdain is again a kind of improving myself or "not being like that" motivation. That's one pair. The second kind of pair is regret and pride, these are first-personal. These are universal sentiments. Regret, for example, its motivational impulse goes to trying to change my judgmental policies in the future. So I regret something where I made a poor decision or made a bad judgment.

**What's different between those two pairs is the first, admiration and disdain, seems to be other-regarding, although it could be self-regarding, too; whereas the other one seems solely self-regarding.**

That's right, that's right. And I think that's the one to focus on, but there are third-personal analogues. I think the first-personal form of it gets at something that I think is really interesting, namely, targeting a particular kind of judgment, a judgment about the worth of various reasons. There are third-person analogues of that, approval and disapproval, which apply to the quality of other peoples' judgments. The third kind of universal sentiment is anger and its pair is gratitude. Each of these is, I think, targeting a different quality of will. The admiration and disdain pair targets quality of character;

> ❝ It seems to me that there are three distinct categories (and these are pairs). The first pair is admiration and disdain. The second kind of pair is regret and pride. The third kind of universal sentiment is *anger* and its pair is *gratitude*. ❞

people's character traits are being evinced by various of their expressions of attitudes and actions.

**Okay, hold on; could you say a little bit more about character because it seems like character is a term that gets thrown around a lot. Again in philosophy of religion I encountered that it's used in a very specific way, so, for example, your character is specifically made up by your free choices. And it wouldn't include things like taste in ice cream or anything like that.**

I don't think it would include taste of ice cream as well but I don't think it's a function of your free choices either. What I'm doing is taking seriously the emotional responses and once I've defined the relevant syndrome, it's to other agents and it has a particular motivation profile, and then seeing what types of thing in an agent are the targets of the emotional response in question. And while we may well admire people plenty for their free choices, whatever those are, I think we also admire or disdain people in virtue of things that are expressions or things they care about that don't have any voluntary component at all. That is to say, things to which they have emotional responses where there's no voluntary component at all. I recall going to a theater once in Arkansas and I was seeing the movie What's Love Got to Do With It, which was about the Ike and Tina Turner story. There were two other people in the theater, who weren't Hell's Angels, but they were from a motorcycle club and they had it on the backs of their jackets, and whenever Ike beat up Tina in that movie, they would just laugh hysterically. Now there's nothing that seems voluntary about that response, but nevertheless, it's absolutely an evocation of a character trait. It's those kinds of things that I want to include under the rubric of things that are admirable or disdainable, contemptible.

**And, presumably, irrespective of their cause?**

I see what you're suggesting. Yes, although this is tricky. So you might have in mind a kind of manipulation case where I'm implant in you the sudden urge to laugh…

**Or just a rotten social background.**

Those are two different kinds of cases; if it's an immediate manipulation that may not be enough to establish the kind of character trait that I have in mind. I'm thinking that character traits are generally persistent (this goes to the identity issue); but to the extent that they have been a persistent part of your life, then, yes, it doesn't matter if they come from a poor formative circumstance or great formative circumstances.

**So you can say, "It's not his fault he's an asshole, but he is a colossal asshole…"**

That's it, that's right; and this is true—there are various predications we make of people, asshole is one of them and that can be absolutely true of somebody.

**So character is to do with your responses, but it's different from your judgment? Explain that distinction.**

When we think of, say, approval or disapproval of what somebody else did, or in the first person case which gets at it more clearly, regret and pride, over what I've done, what am I targeting and regretting? As I said before, it's a kind of poor judgment. This

> ❝ There are various predications we make of people, asshole is one of them and that can be absolutely true of somebody. ❞

may have nothing to do with my character, it's just at that particular time I made a stupid choice and so I regret that, despite the fact that I generally have a very good judgmental character. What precisely is being targeted there? Well I make judgments about the worth of various reasons, and not just the worth of various reasons standing apart, but the worth of various reasons relative to other kinds of reasons. When I make good judgments I have a kind of pride; when I make poor judgments there's a regret, and that's defined by my motivational impulse to try to improve my judgmental policies in the future, that's why I think when we look at that kind of motivational tug, we're seeing what the target of the emotional response is.

**But if you draw too much of a distinction between character and judgment, I'm not sure why I should regret my poor judgment because let's say, I'm going to assume (you can correct me on this or you can argue why you don't think that, but) that my character in some sense is the real me. Then I made a poor judgment, but it wasn't reflective of my character, it's free-floating or something. Unless there's some kind of implication that if I have poor judgment that reflects poorly on my character, then why should I regret my bad judgment? It's almost like something that just happened to me.**

This occupies that middle ground between something that's happened to me and something that is, in fact, reflective on me in a way that might warrant something like shame. To make the distinctions plausible there's got to be these pure cases. A lot of times, of course, all of these responses are going to be overlapping in a single way that my poor judgment reflects on my character in a way that's also reflects poor regard, which is the third kind. In the pure kinds of cases, I think, you can make a poor judgment that is regret-worthy because it's very specific to some context and that it's not reflective of my character overall, but nevertheless is not something that happened to me. These might be cases, say I come out of the Grateful Dead show and I'm just overwhelmed by the love or whatever is going on (I've never been to a Grateful Dead show, but I'm just imagining), and so I come out and I hand all my money to the guy that needs a ride or something like that. Then I find that I don't have a ride home, I don't have a way to get home, so I sheepishly call my wife and I say, ok here's what I just did…

**And she says, "Not again!"**

If she said that, then that would be grounds for a kind of character attribution. But I'm normally very responsible and my mood was heightened in a way, but I wasn't incapacitated. There's something that I think I can appropriately regret, but it doesn't implicate my character so I wouldn't be the appropriate subject of admiration or disdain, but nevertheless disapproval on the part of my wife, regret on my part.

**It's like when you say, "Well that was a stupid thing to do," to someone and he says, "Are you calling me stupid?" And you say, "No—but you did something stupid".**

Yes, that's exactly the distinction I'm trying to capture.

**The first feature, character, is the basis of attributability, and then judgment is behind answerability, which makes sense because judgment has to do with reasons and you can answer to someone on the basis of reasons. Then there's this third feature…**

The third feature is what most people have focused on in talking about responsibility and it's what I call accountability. It's what many people have called accountability as well. I wanted to preserve that label.

**It is unfortunate that there are three As; I kind of wish it was "character responsibility", "judgment responsibility", and "regard responsibility" would have been so much easier.**

I completely agree, and I wanted to preserve the labels that had been used before turns out they're all As, and then I thought hey how cool a triple-A theory, and then I realized very quickly after that this is going to be incredibly confusing. Accountability is just regard, it implicates the quality of your regard for other people, other agents. The emotions, the sentiments that are implicated here are anger and gratitude, that's the universal pair; and anger targets poor quality of regard, gratitude targets good quality of regard. What does that mean, "quality of regard"? It's not judgment, it's not character; instead, it's a particular kind of perceptual stance that you take or have toward other people. And really what it indicates is whether or not you're taking people seriously. It's somewhat hard to explain. There are two kinds of regard: there's evaluational regard and emotional regard. I'll try to illustrate both with examples. Evaluational regard: suppose that I'm thinking about jackhammering my driveway on a Sunday morning at five AM so I can finish the whole project in one day. My neighbor I know works late as a bartender, gets home at three o'clock. There are those facts that are in the air. I'm trying to decide whether to start at five AM. My quality of regard evalutionally depends on how seriously I take my neighbor's interests, given that they will be affected by my jackhammering. The way we typically enable evaluational regard is via empathy. I take up his perspective, I think okay, he's getting in bed and he wakes up at five AM and hears this horrible jackhammering and is just absolutely outraged by it. It's because it matters to him to be able to sleep in on a Sunday morning. (This is going to sound insane, that it can't possibly be the way we do things, but I'm spelling out what I think is a very compressed and quick process and is typically often automatic for well-versed moral agents. I want to pick apart all of the features here.) So I see from his perspective that he assigns worth to being able to sleep in, and to the extent that I can see that, and I return to my own perspective, that becomes for me reason-ish. What I mean by that is a putative reason, the fact that this really matters to him; it's among his interests that he be able to sleep in. To take him seriously is to take that reason as fact seriously and so I may judge then that that fact is in fact a reason that I need to take more seriously and then I may judge that it is serious enough for me to put off the project for a few hours on Sunday morning. That's evaluational regard: it's to be able to see someone else's interest, things they find worthwhile as reason-ish in my own deliberations. The second one is much more straightforward: my wife comes home from work, she's been emotionally traumatized by her boss, let's say, and so she's telling me this story, and she's emotionally wrought and I don't respond with any kind of emotional response, I say "There there dear," I say the sort of right things but I don't respond in a simpatico fashion with her emotionally. I'm a fitting target of a sort of anger in that case, and that's because I failed to emotionally regard her. The two kinds of regard, evaluational regard more colloquially we think of failures of that sort as being inconsiderate. It is failing to consider the interest of somebody else. Emotional regard is insensitivity.

> ❝ There are two kinds of regard: there's *evaluational* regard and *emotional* regard. Evaluational regard: it's to be able to see someone else's interest, things they find worthwhile as reason-ish in my own deliberations. The second one is much more straightforward. Emotional regard is insensitivity. ❞

**In the boss case: suppose you're torn, suppose you are also a boss and you can kind of see his or her point of view, let's say it's a him in this case. You can kind of see his point of view and that's sort of why you're not responding in the way that she wants. Now, are you wrong not to override that, are you wrong not to just have this automatic response of getting enraged for her? In other words, is it your failure to put yourself entirely in her shoes? I think that happens a lot when people say, "Actually I kind of see his point…" That's going to make the person you're talking to even more angry when you get all 'reasonable' with them.**

Unfortunately this is not a work in normative ethics, so I have no answer for you. But what I'm trying to do is figure out what is going on when she responds with anger here and to the extent that we think that it's fitting, then what has to be targeted is my emotional insensitivity. It may be that that is tempered by my taking up the perspective of her boss, and these are the kinds of things that we're negotiating when we talk to one another or we get angry with one another where we think that the anger is unreasonable. But it may also be a function of my ability to step into the perspective of the boss' shoes evalutionally. And here we've got a really interesting case here my evaluational regard for her boss is in tention with my emotional regard for her, and that's the kind of torn that I'm feeling because my regard can split in a couple of different directions.

**It becomes a question of loyalty then. It seems like your empathy shouldn't be applied to her boss because all your empathy should be applied to her and you're being disloyal in taking his point of view.**

Yes and this is the way in which I think the tripartite theory, if it's right, can capture all the nuances of these deeply interpersonal exchanges because what she may realize is that my failure to be emotionally simpatico with her is perfectly appropriate with respect to her emotional trauma, but it reveals a character trait such that she may have contempt for me, or disdain for me in virtue of the fact that I've reflected some objectionable character trait. So there is this kind of ambivalence in these kind of interpersonal exchanges, where the multiple types of responsibility are at war with one another.

**Is all of this making responsibility too broad? This goes back to your "You can't spell responsibility without response," when you have a tripartite theory of responsibility, is this broader than, say, many people mean by responsibility?**

It is, and so there is that battle to fight. Most people just want to talk about what I'm calling accountability, and what they call accountability as well, which is just really about punishment, anger, retributive kinds of emotions. There are a couple of arguments that I want to use to push people in my direction. One is that there are plenty of really smart people—Tim Scanlon and Angela Smith, Pamela Hieronymi, and Matt Talbert, who have pressed for answerability as the only kind of responsibility, and the relevant sorts of responses that we have are much wider than just resentment, indignation, and guilt. For example, for Tim Scanlon, when we blame someone, yes, among those responses might be anger, resentment, and so forth; but also among those responses might be just forming the intention not to deal with this person anymore. Or forming the intention not to deal with this person in matters related to the thing that they wronged you

on. Somebody breaks a promise to you, you decide I'm not going to take seriously the promises he makes anymore or I'm not going to deal with in this particular zone. That seems right to me, that there are ways in which we blame people that don't necessarily include the resentment, indignation, guilt triumvirate. As I say there are these people who focused on answerability and the kinds of broader range of responses there, and also Gary Watson has written a very influential piece called "The Two Faces of Responsibility," in which he focuses not just on accountability but on what he calls attributability which has to do with these kinds of aretaic responses that we've got. There are some people, Julia Driver is one, who say the kinds of aretaic responses you're appealing to—admiration and disdain—yeah, sure, they go to something we'll call moral appraisability, which is what Hume called it, but doesn't go to accountability, so not genuine responsibility. My thought there is: why? It seems to me arbitrary to restrict the range of our responsibility responses just to that small group. Jay Wallace even further restricts them, just to indignation and resentment. And doesn't even include the positive kinds of accountability responses like gratitude, and again this seems to me to be completely arbitrary. What reason do we have to exclude these as among the responses that we have to other people's expressions of practical agency? I'm urging that we take seriously the broader range, and I think that there's not a reason to restrict them. Some people will want to respond to that and say, well look there's some of these responses that implicate things that are not up to us in a way that we think a core of responsibility is about. I have a variety of things about that, which we can talk about.

**Your book is called Responsibility from the Margins, and as you say, you think that it's the key to unlocking responsibility is to look at marginal cases, cases where we feel ambivalence; the psychopath, mildly intellectually disabled individuals, autistic individuals, that we feel ambivalent responses to them, and you want to analyze the ambivalence as, maybe we're not getting all of the three but we are getting one of them. We want to say that this kind of response is allowable and defensible, even if the other two are not defensible. So for the psychopath, in a talk that you gave at UM–Flint, was "From anger to disdain". You say anger is not justifiable, that implies accountability and the psychopath is not accountable because of their various deficits; but disdain is appropriate because their character genuinely is flawed and this is our response to their character. Perhaps you could say why is it that anger is not appropriate for a psychopath, but disdain is? So: say a little bit about psychopaths.**

I'll say a little bit about them after saying a little bit more about accountability. The way we get to genuine regard for other people is: it's enabled by empathy. There's evaluational empathy and emotional empathy and those tracks the two kinds of regard. I think that psychopaths are not accountable in virtue of the fact that they have severe deficiencies if not complete disabilities in their empathic capacities. Their emotional incapacities are fairly well known. Very dulled for anything other than the most basic emotional responses; they tend not to care about much of anything, that they could take up your emotional perspective and feel alongside you in a simpatico way is quite unlikely. I should preface all of these marginal cases by saying that all of them are on a spectrum, and so here I'm talking about a kind of paradigm case of the psychopath who scores over a thirty on the Hare checklist of psychopathy. If somebody is genuinely psychopathic, a full-fledged, is not going to feel emotional empathy. Whether they can feel evaluational empathy is an interesting question but also I think the answer is no, whether or not they can take up your perspective is not at question; they in fact can

> **"** There are some people, Julia Driver is one, who say the kinds of aretaic responses you're appealing to—admiration and disdain—yeah, sure, they go to something we'll call moral appraisability, which is what Hume called it, but doesn't go to accountability, so not genuine responsibility. My thought there is: why? It seems to me arbitrary to restrict the range of our responsibility responses just to that small group. Jay Wallace even further restricts them, just to indignation and resentment. And doesn't even include the positive kinds of accountability responses like gratitude, and again this seems to me to be completely arbitrary. What reason do we have to exclude these as among the responses that we have to other people's expressions of practical agency? **"**

take up other people's perspectives. They've got a theory of mind, they are able to understand what other people are going through, but this is why there's an additional step required for evaluational regard, which is to be able to see others' interests as worthy of pursuit from their perspective, and I think it's that that psychopaths can't see. And the reason they can't see it is that they can't see worth generally. This comes out in their rampant prudential impairments. They pursue one thing and then another they flip from job to job, they make a fortune and they lose it, they make a fortune and then they lose it, those who aren't in prison; they go to prison, they get out; they go to prison, they get out. And what I think this strongly suggests, I think is that they're incapable of seeing worth in their own pursuits, they see no worth anywhere. To the extent that they can't judge that other's pursuits have worth either, then they're incapable of the kind of evaluational regard that's required in my view for taking other people sufficiently seriously. If you can't take someone else sufficiently seriously, then anger is just not appropriate. This is reflected in the kind of pointlessness that you might feel in expressing your anger to a psychopath; it's puzzling, it's amusing to them, but it's not something that they can take seriously.

**So they never get angry at themselves, presumably.**

There is a kind of anger that they feel, but no they don't get angry at themselves from what I understand. They can get angry at other people, and they do have flashes of anger at other people; so something matters to them, but they only get angry when other people are not taking them seriously in a way, or as seriously as they think they ought to be taken. They have very strong desires to do things and if they feel that people are blocking them from being able to do those thing then they have a kind of anger. It's not the kind of anger I think that we have towards other agents, so I mistakenly just said, when they think others aren't taking them seriously, I don't think that the right way to describe their kind of anger. I think their kind of anger is towards others as frustrators of their pursuit of goals. It's not towards others as disrespectors of them.

**But the biggest frustrator of their goals is them!**

Yes, that's true, and they don't get angry at themselves for frustrating their own goals, that's right. I need to be very precise in my language here and I've probably been a little too sloppy. If we think that goals involve a kind of assessment of worth, then I don't want to use that term. I don't want to at all deny that psychopaths have desires, and that desires shape their motivations; it's just that I don't see them as viewing those desires as having any worth in the way that we would judge a pursuit or a goal to have a kind of worth. Psychopaths don't have ends, that's maybe the best way to put that; psychopaths don't have ends. Nevertheless they do get angry when you block the pursuit of their desire, even thought they don't judge it to be worthy of pursuit necessarily. They don't have an evaluative stance toward anything (and again: this is the high-end of the psychopaths). But nevertheless they care about things; they care about the satisfaction of the desires that they happen to have. The fact that they care about the satisfaction of those desires I think is what explains their anger when those desires are thwarted. But when it comes to themselves, it's just they feel whatever the feel at any particular time, if they have a desire to do this, then they do it. And if nothing is getting in their way then there's no reason to get angry at about it but given that there's no long term end that they've got, that they've judged worthy of pursuit, they're

> **"** If somebody is genuinely psychopathic, they in fact can take up other people's perspectives. They've got a theory of mind, they are able to understand what other people are going through, but this is why there's an additional step required for evaluational regard, which is to be able to see others' interests as worthy of pursuit from their perspective, and I think it's that that psychopaths can't see. And the reason they can't see it is that they can't see worth generally. **"**

not going to get angry at themselves for frustrating that because they don't have such things.

**Perhaps their problem is that they're extreme "episodics," to use Galen Strawson's term. Maybe you could tie this into personal identity; as you say, they seem to get very invested in certain careers, they can be very successful salesmen, you point to a specific example about a guy who made pots of money, but then—bam—it's gone. If you ask them later, "Well, do you regret screwing that up?" do you know what they would say? Would they say, "No, I have no regrets?"**

Yes, there are no regrets.

**Then one way to understand that is to say they don't invest in that former self; that they don't identify with that former self.**

That's a really interesting point, there's a piece by Gary Watson, "Psychopathic Agency and Prudential Deficits" (Proceedings of the Aristotelian Society 113, October 2013), where he's arguing that psychopaths lack a kind of practical identity. This is really interesting, I'm wondering if Galen Strawson would, because he thinks he's an episodic, whether he would reject the notion that he lacks a practical identity if he's genuinely an episodic. Maybe he rejects the coherence of a notion of a practical identity, I'm not sure, but this is a really interesting project to pursue because it has been thought by some people that psychopaths are really, what Harry Frankfurt called a "wanton;" they don't care about themselves, they don't care about anything, they just pursue whatever it is they pursue at the first order desire level, they have no second order reflective desires about those first order pursuits. That does seem somewhat right to me, although I do want to resist the notion that they care about nothing; I think they care about at least the satisfaction of whatever desires they have at any individual moment, given that they have such strong anger in response to the frustrations of their desires.

**Do you think that it could turn out that "psychopath" is like "neurotic"? It's like a term that will be abandoned when we get better science? Do you think neuroscience could show that, in fact, there aren't really psychopaths as we understand them, or are you fairly convinced that psychopath is a natural kind?**

I think what's going to happen is we're going to see a separation in the psychological literature about it, because there are two factors that are relevant to the diagnosis of psychopathy. Originally it was put into place as something to diagnose in trying to identify what it is that led people to be recidivists as criminals. If a person scored highly on this test, it was going to be able to predict whether or not they were going to be recidivists. But there are two factors: one is in the Hare checklist, a serious of behavioral elements; the other is a set of serious, primarily psychological elements, and the latter includes things like lack of empathy, callousness, lack of remorse, lack of regret, these sorts of things. I think what's going to happen is that the anti-social personality side of it, which is what explains why lots of folks are in prison is going to be the focus and that will split off from the factor one features of psychopathy, which are the more psychological features, and the ones—these are the kinds of individuals that philosophers, for example, are more interested in. Those who are very high on the factor one scores (callous, lack of emotion), but don't have the behavioral issues tend to

> **❝** It has been thought by some people that psychopaths don't care about themselves, they don't care about anything, they just pursue whatever it is they pursue at the first order desire level, they have no second order reflective desires about those first order pursuits. That does seem somewhat right to me, although I do want to resist the notion that they care about nothing; I think they care about at least the satisfaction of whatever desires they have at any individual moment, given that they have such strong anger in response to the frustrations of their desires. **❞**

be "successful" psychopaths, and so are out and about in the world, not in prison—yet, anyway. I think what's going to happen is that they will split, but in that split it's still going to be a tricky issue whether the factor-one psychopaths are going to constitute a natural kind, which was your original question because it's not entirely clear what kind of person that would predict. I'm not sure, I think the jury is still out on whether psychopathy is going to last or if will become, like you said, "the neurotic".

**I'm wondering if you think it will map onto a part of the brain? Or even if there is no part of the brain that it maps onto, it's a recognizable enough deficit that we can pick it out some other way?**

I think it's a recognizable enough of a deficit that it's very likely that there's a particular part of the brain it maps onto and I think, work like that of James Blair is honing in on those particular neuro-biological features, that there does seem to be relevant sort of deficit that might be able to, in a unified way anyway, explain many of the factor-one features—the lack of empathy, the callousness, the shallow emotional affect, the impairments in responding to fear (they do skin conductance tests and the things that would frighten most people there's no real response that they have, or very limited response)—looking for ways in which all of these impairments, disabilities are probably going to map onto particular location, or small subset of locations.

**What's the difference between psychopath and sociopath? Is there one?**

It's not really used, most people have gravitated toward psychopath. It was never entirely clear what the distinction was supposed to be anyway and at least for psychologists, who are using the term, despite the fact that it's not in the DSM-5 yet, almost invariably the term is psychopath, so that's what most people are going with now.

**Tell us about NOWAR [New Orleans Workshop on Agency and Responsibility].**

This fall will be the third. We put out an invitation for abstracts for papers in the November of the previous year and then usually get about 150 abstracts and from that nine are chosen to present. We have two keynote speakers this all, Julia Driver and Derk Pereboom, and the theme is anything pertaining to agency and responsibility and so we're really looking for people across disciplines that are interested in these issues, and then the papers are collected every other year and published in the Oxford Studies in Agency and Responsibility.

**The last time was the one on "Freedom and Resentment," or the time before?**

That turned out to be a supplemental volume we did in between the normal biannual NOWAR; that was a conference that they did at William and Mary that Neil Tognazzini and Justin Coates put together on the fiftieth anniversary of the article "Freedom and Resentment"; so we collected those papers and put them out. The second official volume, the third volume really of Oxford Studies should be out any day now, it's from the conference here in New Orleans in 2013.

> 66 Those who are very high on the factor one scores (callous, lack of emotion), but don't have the behavioral issues tend to be "successful" psychopaths, and so are out and about in the world, not in prison—yet, anyway. I think what's going to happen is that they will split, but in that split it's still going to be a tricky issue whether the factor-one psychopaths are going to constitute a natural kind. 99

**Kantians, and Kant in particular (if he is a Kantian) are notorious for saying that you have to be an agent to be a person, or to be worthy of moral regard and we can be nice to animals and children, but we're basically doing them a favor or maybe (for children) because they will become fully-fledged agents. If we take your tripartite theory as saying something like you have to have these three features to a sufficient degree to be capable of fully fledged responsibility, let's say that's what it takes to be an agent. Does that mean someone like a psychopath in some respects doesn't deserve moral regard, maybe we should say they lack certain rights, or we don't have to regard them as morally on a par with regular folks who meet all three criteria?**

That's an interesting question, and I don't think that's the case; it may be that certain entities are of course owed moral regard independently of whether they're capable of showing us any regard, and I think that's surely true of psychopaths and maybe some animals and children and so forth. I wouldn't want to rest the regard that we show for others on their capacity regard for us.

**What do you think, given how you were drawn away from an actually contributing career to become a philosopher, what do you think philosophy's role is in the public sphere?**

There are a variety of roles and this may be a strange analogy at first, but bear with me. Philosophers are in a sense like fashion designers. There are varieties of kinds of fashion designers; there are those who are designers for what people wear everyday and you go all the way up to haute couture, where no one wears that outrageous stuff. But nevertheless, the kinds of ideas that are being put forth in couture gradually make their way down in surprising and interesting ways into the things that people wear every day. So I think there are a variety of roles for philosophers that mirror that aesthetic structure. There are people who are doing haute couture philosophy; purely theoretical that in its way may filter down to more everyday concerns, where other people obviously doing on-the-ground applied ethics. The role that philosophers ought to play with respect to the public, I think, is going to depend on what interests that they have, that's not to say they won't every contribute to the public even if they're doing that kind of pure theoretical stuff. But I think that we, especially those of us who are doing ethics, even if it's of a more theoretical bent, do need to speak out more because we have things to say in the public eye. This is happening and I'm very exited to see this happening in some of the articles in The New York Times, "The Stone," and some of their op-ed pages. These things have been occurring in a way that's exciting to me, or the various blogs that have gotten some public traction. Part of the issue is that there are certain areas in applied ethics, for example, which would be the natural zone for philosophers to contribute the most, where I think philosophers themselves view the issues in question as so settled that there's no need for us to talk about it. That's just false for the general public and I think that they could be of great value in contributing.

**For example?**

I was thinking in particular of something that Jeff McMahan said about gun control. I think amongst philosophers, for example, this is thought to be fairly settled, of course there should be serious restrictions in who should own guns and there are very good reasons for it that can be given, but as I say lots of philosophers think, well there's no

> ❝ Philosophers are in a sense like fashion designers. There are varieties of kinds of fashion designers; there are those who are designers for what people wear everyday and you go all the way up to haute couture, where no one wears that outrageous stuff. But nevertheless, the kinds of ideas that are being put forth in couture gradually make their way down in surprising and interesting ways into the things that people wear every day. So I think there are a variety of roles for philosophers that mirror that aesthetic structure. There are people who are doing haute couture philosophy; purely theoretical that in its way may filter down to more everyday concerns, where other people obviously doing on-the-ground applied ethics. ❞

reason to talk about this, it seems like a dead end issue. Or, until very recently, same-sex marriage was thought to be such an obviously supportable right thing to do that amongst philosophers that people had stopped teaching it in their courses for the last five years, people haven't even thought it anymore because they've thought there just aren't any good arguments on the opposing side. Well, bring those to bear to the public and I think that can be a very valuable service. Talk about why you think the opposing arguments are bad arguments.

**Don't you sometimes feel though that arguments are never going do anything? Gay marriage is one of those things that amazes people of our generation because it happened so fast. When I was in graduate school, when I was a TA, the idea of gay marriage becoming legal within twenty or thirty years would have been bewildering but I think it's more to do with social media, reality TV, stuff like that. For example, on the one hand it seems kind of obnoxious that Caitlyn Jenner is getting all this press because haven't we had enough of the Kardashians, now we have to respect one of them? But that's how progress gets made, in these ways that, to a philosopher seem infuriating because it's like, "But we've been talking about this for centuries and suddenly you catch on because it affects one of the Kardashians?"**

You're exactly right, and so this is where philosophers are not fully aware or take sufficient consideration of the way in which social change is affected. A social change is not generated typically by argument and argument alone, and this is maddening to people. On the same-sex issue, for example, what finally changed people's minds was more people coming out who were their family members, and you get Republican politicians saying something like, "Oh I was against it until I saw that my son or son-in-law was gay." What I'm trying to do in the classroom is to teach people not to rely on that kind of change in attitude, but to be able to assess the arguments for things independently of the attitudes we might have, independently of those affections we might or might not have because you may not know anybody who is of an oppressed minority but nevertheless you ought to be able to recognize that they ought not be oppressed. I've done a kind of change in my own view on this in the last year or so, when it comes to the reasons that we think are relevant or irrelevant, that's not the way that public policy change is generally affected. In order to have more of a voice in the public square we have to recognize change is incremental typically, although the same sex case is a radical exception to that; as you say it blew my mind. Most change is fairly incremental, however bizarrely some policy has come into effect, as most of our policies are, it's typically the result of crazy history and negotiation between parties and legislatures and so forth. We have to recognize that that's the case, for example, that there are no good reasons to differentiate drug use from alcohol use, for one to be legal and one isn't. Nevertheless that's the case and so if you want to change drug policy you have to start with that recognition you can't say, look there's no distinction between alcohol and drugs. For a variety of crazy reasons drugs have become prohibited and you have to start from that fact and try to move from there if you're going to go for a more open policy. It's a very different set of reasons, I think, that is relevant in the public square. Philosophers typically don't want to engage with those kinds of reasons because they don't think they're relevant and if we are to affect social policy change we have to recognize this.

> 66 I've done a kind of change in my own view on this in the last year or so, when it comes to the reasons that we think are relevant or irrelevant, that's not the way that public policy change is generally affected. It's a very different set of reasons, I think, that is relevant in the public square. Philosophers typically don't want to engage with those kinds of reasons because they don't think they're relevant and if we are to affect social policy change we have to recognize this. 99

**Are there any radically different areas of philosophy you intend to explore, is there something that when you're teaching Philosophy 101 you say I've always thought this was cool and I've never had a chance really to explore it?**

I've been interested in philosophy of religion for years and haven't really written anything about it.

**Now there's an area where you actually get intellectual figures, becoming famous, sometimes the wrong ones like Sam Harris, but still there's a debate that seems to be happening or at least has been happening over the past decade.**

That's absolutely true and I've been somewhat encouraged and somewhat discouraged by the tenor of that debate, but nevertheless that's something that might be worth pursuing.

**You kind of see what happens, though, to these intellectuals that enter the public sphere, it suddenly becomes about their personality. Even someone like a lefty-leaning website like Salon.com runs all these salacious things about the New Atheists don't understand about religion and things like that. You can't make it just about issues; you have to be prepared, I guess, if you're going to be a public intellectual.**

That's right. Some of them are doing a decent job at it, but you're taking quite a bit on when you do that and you'd have to prepare yourself well.

**And watch what you say, which Dawkins doesn't seem very good at.**

**There's a ton more to be said about the subject of your new book (Responsibility from the Margins) but you are limited in time and the reliability of your power source, so—thank you very much!**