

Dear Conference participants.

As part of our presentation at the conference, Ken Kendler and I thought it would be interesting to provide some discussion of our process of collaborating on two papers that we have written together. This will be in addition to a quick discussion of the contents of the papers themselves. One of the papers is on the use of polygenic risk scores and is currently under submission. I will present some of the ideas in this paper and, time permitting, will also reflect on some of the themes about collaboration immediately below. Ken will present on our paper on top down causation in psychiatric disorders, a copy of which will also be posted to the conference website. Ken will add his reflections on collaboration to the website by this coming Saturday.

Jim Woodward

### **Woodward presentation on writing the PRS paper.**

After Ken and I collaborated on the top-down causation paper, as he has described, we decided it would be fun and satisfying to work on another collaboration. I recall several conversations about a possible topic. I believe I originally suggested something on GxE correlations but, as we talked, it became clear that, as interesting and important as this was, it was already the topic of a huge amount of scientific discussion and it wasn't clear that we would have anything new to say-- or at least it wasn't clear to me that *I* had anything new to say. Then Ken suggested polygenic risk scores, which he had used in his own work and were a topic of considerable current interest, both in connection with mental illness and in social and behavioral science applications. I knew virtually nothing about these scores-- I think my only previous acquaintance with them was via a conversation with Eric Turkheimer (who I knew previously and was present at one of those Copenhagen meetings Ken talked about) in which he was quite critical of their use.

Nonetheless I was intrigued and Ken and I decided to write about this topic. I believe the first step we took involved Ken sending me a bunch of papers of varying levels of technicality, both applying PRSs and discussing general features of their construction. I remember finding some of these papers quite challenging-- in part because I was unfamiliar with a lot of the vocabulary that was used. (In some cases, it turned out I was familiar with the underlying concepts or at least found them straightforward enough-- it was just the words that were unfamiliar-- e.g., use of distinctive biomedical language for fairly familiar statistical ideas. In other cases I had to dive into genetics textbooks, Wikipedia articles etc. to understand) In any case, this resulted in a lot of emails full of questions from me to Ken which he answered very patiently. I believe that we also had several long zooms sorting out various confusions I had.

I've had a long standing interest in the social and behavioral sciences so one of the things I also started doing was reading various high profile papers using the PRS methodology to help understand various sorts of socially important outcomes, such as educational attainment-- this was a bit in contrast with Ken's interests which were mainly in mental disorders. I think it was

about this time that I decided to reach out to Kathryn Harden who was one of the primary researchers using this methodology in social science and educational applications and ask if I could schedule a zoom with her, because I found some of the claims in her papers rather puzzling. (We had a mutual acquaintance who had told me about some of her work earlier and he helped arrange this.) It was also around this time that I learned about Harden's forthcoming, now published book, *The Genetic Lottery*, which has attracted a lot of attention. Coincidentally about this time I was asked by a psychology journal to review a very interesting paper by Harden and a grad student of hers, James Madole, about the interpretation of PRS and in particular about the extent to which these could be given a causal interpretation via a supposed analogy between the random nature of meiosis and randomized controlled trials. So I ended up having a long zoom session with Harden and Madole as well as a number of follow up emails which greatly helped to clarify for me what some of the most interesting issues around PRSs were.

I also did a dive into the recent philosophical literature on genetic causation, the interpretation of SNP/trait correlations, GWAS and so on. I was previously familiar with some of this work but by no means all of the recent stuff. I was, frankly, pretty unhappy with what I found-- lots of incorrect empirical claims, a high level of naivete about the conditions that need to be satisfied for reliable causal inference from observational data, misunderstandings about probability etc-- e.g., the amazing claim that the laws of large numbers somehow ensure that correct causal conclusions regarding gene/trait relations will emerge from purely correlational data (e.g., from a GWAS), as long as there is enough of it, without any need for additional assumptions. This presented a dilemma. On the one hand, I knew that we had to make some contact with this literature since we planned on submitting to a philosophy of science journal. On the other hand, I didn't want to use up a lot of space in the paper criticizing the extant literature, which was just going to distract from what I thought was most interesting about PRSs.

So I basically proceeded by reading pretty much anything that I could find regarding PRSs -- did this for several months in fall of 2020, making notes but not trying to do any writing (because I thought there was too much I did not understand) and talking to people like Ken and Harden. I sent Ken an outline for a possible paper in early January, 2021, and we had zoom and email discussions around that. Then I started writing. Our joint paper went through many, many drafts-- I haven't tried to count them but easily more than a dozen. Ken corrected a number of scientific mistakes and infelicities in what I wrote and also added scientific background and context, including connections with the Fisherian tradition which I certainly would not have understood on my own. Making decisions about what should go into the paper was not straightforward since there was so much that was interesting and novel about PRSs. Partly because of my interests I found myself drawn to issues about the causal interpretation of PRSs and their potential use in causal inference while I think that Ken's primary interests were somewhat different. I think that we ended up focusing to some extent on issues of causal interpretation in the paper because we needed organizing themes and also because that (the causal stuff) was one of the main preoccupations of the related philosophical literature.

I wish I could report that the paper had smooth sailing after we finally completed a version we were ready to submit but this was not the case. The first (philosophy of science) journal to which we submitted rejected it, citing among other defects insufficient engagement with the philosophical literature. One reviewer was particularly unhappy with the absence of a sufficiently detailed and appreciative discussion of the views of a certain philosopher (almost

certainly the reviewer). The other reviewer complained that our paper lacked a proper introduction for philosophical readers and that the content of the paper was "meager" but without providing any further detail. Neither reviewer provided any objections either to empirical or conceptual/methodological claims we made in the paper--- the view of the reviewers seemed to be that the paper should be rejected because it did not "look" like a standard philosophy paper-- at least one of the reviewers said this explicitly. (I think that whatever else is true of the paper, it certainly has plenty of philosophy in it, even if it did not conform to the reviewers stereotype of what a philosophy paper should look like.) The paper is now under submission at another journal, so we will see.

### **Email exchange between Ken and me that helped to generate some of our reflections about collaboration.**

Ken: this is from an email exchange we had, which you suggested might provide one theme for our talk. One possibility is to present this (as illustrating some of our frustrations) to the audience. I could then follow with the reflections below.

[ From Jim] Ken, I completely agree with the sentiments you express below. It is an unfortunately common experience for me to listen to some philosopher talking about a topic in the philosophy of X where X might be psychiatry, philosophy, biology, economics, even physics and where it is clear to me, even as a non-expert in any of these disciplines, that the philosopher has some basic misunderstandings of either empirical or conceptual/mathematical issues in the discipline. Even when the philosopher has a good graduate school text book level of understanding, lots of subtleties about the way in which insiders in the discipline think are missed— it is though the philosopher has acquired the letter but not the spirit. The result is a lot of wasted time and energy. The obvious solution seems to me to collaborate with someone who knows what they are talking about in discipline X. By no means all philosophers who want to do philosophy of X are willing to do this, however.

[From Jim] ... one of my messages [ if you want to do philosophy of X] would be “find someone who works directly in X and collaborate with that person”. Might seem like common sense but by no means always followed in philosophy.

[From Ken]

I feel the same way from the “other side of the fence.” Not infrequently, I get asked by young psychiatrists how do I get started in the philosophy of psychiatry? My response is “unless you have a substantial background in whatever area of philosophy you want to apply (very few do) find an interested expert philosopher who can at least mentor you and maybe, even better, collaborate.” Not many seem to do that. I have heard a lot of talks on the “philosophy of psychiatry” where the psychiatric assumptions on which they are based are painfully inadequate. Our Copenhagen mtgs have the not so hidden agenda of trying to “hook up” philosophers and mental health professionals to possibly collaborate. Ken

[From Ken] We have at least one theme for our talk. I think I told you attending a “philo of psychiatry” symposium at the PhiloSci mtg to which you invited me. I went up afterward to discuss the talk with one speaker with the hope of giving her advice and references about where she was erring in her discussions on a psychiatric topic I knew well. She clearly was not interested. That was really discouraging. Ken

[From Ken] I have one more anecdote on this topic. In one of my first discussions with John Campbell ever, while at the Stanford Ctr, he discussed a theory of delusions in the Capgras Syndrome (Capgras syndrome is characterized by a delusional belief that a person has been replaced by an imposter) resulting from changes in sensory perception. He said this was based on a famous much discussed case. I told him of one of the more striking clinical experiences in my residency of seeing a Capgras syndrome emerge right in front of me in a way that clearly had nothing to do with issues of sensory perceptions. His jaw dropped. He never knew other cases of this, saying that a number of philosophy articles had argued about this theory all based on the one case. No one had, apparently, bothered to ask psychiatrists about other possible cases!

### **General Reflections on Doing Philosophy of X-- Woodward**

As recounted above, my one sentence advice to a philosopher who wants to do philosophy of X (where X may be biology, psychology, physics etc.) is: find someone who works directly in X and collaborate with that person. Failing that, find someone in X who is willing to mentor you, patiently answer your questions, direct you to relevant literature and so on.

I've tried to follow this advice not just in my collaborations with Ken but in a number of other of philosophy of X projects. For example, three of the papers I've been involved in on the neural processing underlying moral decision making have been collaborative efforts with neuroscientists and neuroeconomists and two of these were published in neurobiology/physiology journals. A long empirical paper describing the results of experiments testing some ideas I had about causal cognition in young children was co-authored with a number of developmental psychologists and published in a psychology journal. Similarly for a collaborative effort testing some ideas of mine about the connection between a certain causal concept and responsibility attributions.

This recommendation about finding someone in X to collaborate with is non-trivial at least in the sense that a large number of philosophers who do philosophy of X do not follow it. (It is an interesting empirical question how many philosophers do. My impression is that for many X's the majority of philosophers of X are not involved in serious collaborations with researchers in discipline X, although I may be wrong about that. )

Why do many philosophers of X fail to collaborate with researchers in X? No doubt there are many reasons. Here are a few: 1) The philosophers don't think this is necessary, given their purposes. They figure they can learn enough about X on their own (by reading papers, textbooks in X or perhaps just by reading the work of other philosophers of X. 2) The

philosopher can't find anyone in X who wants to collaborate with them, at least on problems that they (the philosopher) want to work on. 3) Related in part to 1) To get published in a philosophy journal on philosophy of X, it is often sufficient and in many cases absolutely necessary to discuss what other philosophers of X have done, -- there is no corresponding requirement to pay detailed attention to research in X itself, what over all current opinion in X is on the topic discussed etc.

Let me begin with 1). There is a cliché about the philosopher who thinks that all problems in discipline X could be resolved by a smart philosopher (herself or himself of course) who is prepared to devote a Saturday afternoon to the topic. Of course serious philosophers of X are not supposed to think this way but I believe that something of this underlying attitude persists among many of them-- and is part of the reason for the "I'll just learn about it myself" attitude described under 1). The problem with this strategy is that for most philosophers, most of the time, it just doesn't work well. Speaking for myself, I know perfectly well when I try to do something in philosophy of X, even after diligent reading and thinking, that there is a lot that I don't understand-- that is why I feel a need to collaborate.

As suggested by the email exchange between Ken and myself above, I'm afraid that it is a fairly common experience to listen to some philosopher talking about a topic in the philosophy of X, where it is clear (even to me as a non-expert, let alone an expert in X) that the speaker has some basic misunderstandings of either empirical or conceptual/mathematical issues in the discipline. When I check with others, either in the discipline in question or with philosophers who are well-informed about the topic under discussion, they confirm my assessment. In fact, my experience is that if you talk to people in discipline X or the relatively small number of philosophers who are well informed about X, they will tell you (to the extent they are paying attention to it at all) that much of the work done by philosophers of X suffers from obvious empirical and conceptual problems. Even when the philosopher has worked through a lot of graduate school level text books in X, there often will be lots of subtleties about the way in which insiders in the discipline think that are missed. This is particularly likely to be the case if, as often happens, the philosopher brings various philosophical preconceptions to problems in discipline X which are at variance with the way people in X think. All of this suggests to me that collaboration brings lots of benefits that are not otherwise easily attained.

How about 2) -- philosopher has trouble finding someone in X who is interested in the problems the philosopher is interested in? One piece of advice is to try harder to find a collaborator but there is something more to be said. Perhaps your difficulty tells you something about the status of the problems you are interested in-- that researchers working in X don't think it is a real or interesting problem or at least a problem that connects naturally and organically with concerns within discipline X. This is certainly information that the philosopher of X should pay attention to. Of course it can happen that there are genuine problems within X that researchers in X don't recognize as problems but that the philosopher can see *are* problems but I think that one has to know quite a lot about X before one is in a position to make this judgment. Let me add that I have no doubt that in many disciplines there are conceptual/analytic or methodological problems-- problems that are in that sense philosophical -- that arise in a natural way and that sometimes philosophers have skills that can help with solutions with these problems. I'm just suggesting a bit more humility among would-be-philosophers of X about problems in X that are worth paying attention to. Of course it is also completely legitimate and worthwhile for philosophers to operate simply as consumers of results in X, using these to cast

light on some issue, the motivation for which comes mainly from philosophy itself. But if this is what one is doing, it should be clear that this is the goal.

As for 3), there is no doubt that these are unfortunate incentives and the question is how they might best be changed. One suggestion for editors: when a reviewer objects to a paper on the grounds that it does not engage sufficiently with "the literature" on philosophy of X, the editor should ask (and the reviewer should ask him or herself) why this sort of engagement is necessary given the goals of the paper. It should not be presumed that the only legitimate goal of a paper in philosophy of X should be to rehash some of the standard philosophical literature on some topic in X. A somewhat more radical suggestion is this: there ought to be some presumption that those doing philosophy of X should sometimes publish in journals in discipline X and one (not the only but one) criterion for judging work in philosophy of X should be whether researchers in X pay any attention to it (assessed in terms of citations etc.) As philosophers of X we should not be satisfied with discussions of some topic that is conducted entirely within philosophy journals and has no visibility with discipline X itself. A professional norm of this sort would help to avoid situations in which the philosophers of X talk only to one another in philosophy journals.

In saying this, I don't at all mean to suggest that issues that are, so to speak, more purely philosophical or that are more general than those taken up in any particular scientific discipline should be excluded from the purview of philosophers of X. But when such issues are addressed one should be clear headed about what is being done and what the goal is. Using examples from X to motivate general claims about, say, reduction or the nature of scientific representation or the possibility of non-causal forms of explanation can certainly be worthwhile but it is (at least often) different from making a contribution to X itself. Even in this case a reality check from researchers within X concerning one's use and interpretation of examples is often a good idea.