

**GiveWell San Francisco Research Event June 19, 2018 – Open  
Philanthropy Project**

*This transcript was compiled by an outside contractor, and GiveWell did not review it in full before publishing, so it is possible that parts of the audio were inaccurately transcribed. If you have questions about any part of this transcript, please review the original audio recording that was posted along with these notes*

00:00 Holden Karnofsky: So I've hit record now, and as usual, if you want to ask a question and you don't want to record it, or if there's anything else you don't want recorded, just tell me at the time or tell me afterward and we'll make sure to take care of that. I'll also hit record there. I'll be repeating questions for the recording. So yeah, I'm Holden Karnofsky. I'm the co-founder of GiveWell, and currently I'm not at GiveWell, I'm the Executive Director of the Open Philanthropy Project. So, I think most of you are pretty familiar with the similarities and differences. I'll be pretty quick about what we are and mostly focus on what's kind of new and what we're up to these days, but the really quick overview, GiveWell tries to help people give away money as well as possible. When I was still at GiveWell, Elie and I met Cari Tuna and Dustin Moskovitz, who are giving away a very large personal fortune, like in the billions of dollars, over the course of their lifetime, and the Open Philanthropy Project was started as an attempt to help them give away their money as well as possible, which we felt called for a slightly... Or well, as it turned out, kind of completely different approach from GiveWell, where I would say GiveWell is really focused on the most evidence-based giving possible. A lot of what they do is the kind of thing that you can make a case for.

01:14 HK: You can spell out every step of the case, and you can sort of win credibility in a public setting for that kind of giving, and Open Philanthropy is often trying to lean the other way and do the things that only a really major philanthropist is well positioned to do. And so, a lot of times what that means is, taking... Maybe sometimes creating new organizations with money, sometimes completely transforming organizations, and generally doing things that could be extremely hard to measure, where measurement might not work at all, where it's very long-term, very high risk and often requires getting so deep into the issues that you find it's not tractable to explain what you're doing to a non-expert at all. And so, they're kind of... Both organizations are all about doing the most good possible, with giving away money, but we've grown into these almost diametrically opposite approaches. I think we both really try to be all about being analytical about giving, being intellectual about it, trying to do the most good possible. And so, we have that in common, and we're both big fans of evidence and getting educated about the issues we work on. But a lot of times, GiveWell is the stuff that is easiest to make a public case for and Open Philanthropy is the stuff that it's hardest to do that for. So that is the kind of genesis of the organization and the differences between them.

02:31 HK: In terms of where Open Phil is at right now, we spent our first couple of years just picking focus areas, and so, we tried to pick sort of causes focus areas, problems to work on, based on how important, neglected and tractable they are. So, we try to pick areas that affect a lot of persons and affect them a lot. That's important. Areas that aren't getting too much attention from other actors. So there's something different for us to do, and that's neglected. And then, we try to pick areas where we see some kind of path to success, and that's tractability. And so, we spent our first couple of years choosing areas to work in, based on those three criteria and then hiring the people who we thought would best lead our work in those areas. And so, the basic Open Phil model is pick the best causes to work on, pick the best people to work on them. And then, from there, we really try to structure everything we do to give those people freedom to be creative, to do things that are different, to do things that they won't always even be able to explain to my satisfaction, to me, as someone who's paying a lot of attention. And so, a lot of our work is about minimizing veto points, minimizing the need for consensus, maximizing the ability to do things that are hard to explain, and are big risks, and that are bold. So, that's kind of the philosophy, that's how it works.

03:47 HK: In terms of how much we're giving and what causes we work in, currently, we're running at about a bit over \$100 million a year in charitable giving. That does not include the money that Cari and Dustin are giving to GiveWell's top charities, so that is part of our process to determine what that number should be. But when I say over \$100 million, I'm excluding that. Main causes we work on at the moment, we work on criminal justice reform, which is a US policy cause, we feel that incarceration in the US is just at extraordinarily high levels and way beyond what is in the public interest. There's a lot of suffering as a result of that, and we feel that it is possible and important to reduce incarceration while maintaining or improving public safety. That's a cause we picked a lot because of the tractability, because we saw opportunities to actually make a difference there.

04:39 HK: We work on farm animal welfare. So, this is an interesting cause, where we believe there's huge numbers of animals being treated incredibly poorly on factory farms every year. And it's a really open question, like, how much should you value animals' lives, and animals' experiences compared to people? But if you... Even at a number, like valuing them at 1% or something on a year by year basis, the amount of suffering on factory farms is enormous, and the amount of opportunity we've seen to do something about it, is enormous. So, I'll be getting to those a bit later. We have

a major scientific research program, we try to fund neglected and important science that is really promising, that can't get funded through the normal mechanisms. And then we have a really big interest in global catastrophic risks. So one of the ways to help a lot of persons for a little money, one of the ways to do that is to look at possible low probability events that could affect the world so much that they might actually affect large numbers of future generations.

05:42 HK: So an example of this is, if you imagined... If you imagine due to climate change or something that it became impossible to build a civilization on this planet anymore, that wouldn't just... If you could prevent something bad like that from happening, you wouldn't just be helping the seven or eight billion people who are alive today, you could be argued to be helping the many, many future generations of people who ever will be alive. And so, global catastrophic risks are a major interest for us. Certainly, climate change is one of the things we work on. Also, our major priorities there are biosecurity and pandemic preparedness, so when we think about the worst things that could happen, many of them are some sort of pandemic. And then potential risk from advanced artificial intelligence, which is a cause that... Just the basic idea of the cause would take a while to explain. So I'll move on for the moment.

[laughter]

06:30 HK: So our giving is over \$100 million a year right now, and in terms of where we are as an organization, we are doing that, we've been at that level for a couple of years. We're starting to see hints that that giving is having an impact. And so examples of that, especially in the more short-term causes where we pick them partly because we saw opportunities to have wins. So, on criminal justice reform, we've seen a couple of bipartisan bills in Illinois, that we believe, one of our major grantees made a major contribution to it partly because of our funding, and we believe that will have a big impact on the prison population in Illinois for years to come. That's very exciting, a good impact, lower prison population.

07:12 HK: And also we have been very interested in the culture of prosecutors. So public attorneys, where we believe that the default way that prosecutors run for office is they emphasize how tough they are on crime and how many people they are able to lock up. We're interested in an alternative vision of prosecution where a head prosecutor tries to serve their community by weighing cost and benefits and by thinking, "When we lock someone up that's bad, it costs the community money and it costs that person and their family and their friends." And so, maybe the job of a prosecutor is not to lock up the most people, but to do right by their community, do the best things for their community. And we've seen a mild beginning of a wave of prosecutors who think more that way and that's been a major focus of ours. Nothing we do, none of the good things we see in the world are because of Open Phil. We're always supporting grantees and we're always only a part of the picture. But that has been a major focus, and we're starting to see progress.

08:09 HK: Larry Krasner in Philadelphia is probably the most obvious example of a prosecutor who is doing things differently. He put out a memo that said, "When people are asking for a prison sentence, they've got to calculate the cost of that prison sentence." It's like tens of thousands of dollars a year to the government, and they have to argue for why putting this person away in prison is worth all that money and why they couldn't put that person away for less time and save the money. This person is also requiring all prosecutors to get his direct sign-off any time they're plea-bargaining with someone for more than... Starting at more than the minimum sentence as their offer. So just taking a different approach there and has done a whole bunch of other interesting things.

08:49 HK: The farm animal welfare work too, I think we're starting to see hints of impact. So when we got involved in that work, there had been a few big exciting victories, such as getting McDonalds to pledge to go cage-free. So, McDonalds had made a pledge that they would make a move towards selling only eggs that came from a cage-free standard, which we believe is much better for chickens than a normal standard. And we came in, there'd be a few victories like this, and we put money into the major corporate campaigners and saw things really explode. And again, I don't want to say this is because of Open Phil, but we do believe we're part of the picture. We sped things up, we made things more certain, better funded, faster. And since then, essentially every major fast food company and every major grocer in the US has made cage-free pledges now. And so hopefully 10 years or so from now, it will be almost impossible to get eggs in the US that aren't cage-free. And that is an enormous amount of potential suffering averted, and I think it's also good for the long-term health of the movement to be getting some wins.

09:52 HK: We've taken that work, and now we're looking forward and going for the next victory. So looking at the way that boiler chickens are treated in the US, and also looking abroad at whether we can help build similar movements in India, China, Latin America. And try and be part of those earlier victories and not just come in as the wave kind of grows. So that's been exciting as well.

10:16 HK: We're also, right now, we're really focused on helping Cari and Dustin give away their own fortune as effectively as possible, and we're going to need to eventually raise the amount of annual giving to meet their targets. So we're not currently focused on advising other donors, and on raising money from other donors, but nonetheless, we have started to see hints of that as well. So I think also in the two causes I just mentioned, criminal justice and farm animal welfare, those are two causes where there's a lot to do, there's a lot of groups ready to fund, we haven't wanted to scale up our funding too quickly. And so we have given the program officers budgets, and they've ended up supplementing those budgets by raising money from other donors. And that's been interesting to see that they've come to be seen as experts in the field.

11:00 HK: Criminal justice in particular, I think the money raised has been kind of competitive, or a significant chunk, or in the ballpark of the money that we've been providing, and so that's been cool to see. So we're at a pretty high giving level now. We're starting to see some early signs of success and that said, we don't want to significantly grow our giving to its eventual peak until we've done more work on ourselves, in a bunch of ways. So one is just having more time to see how our giving goes. Question our assumptions, which is something we've done a lot of. We've sort of done a bunch of examinations of assumptions, like examining our view that it would be good to reduce the incarceration rate, examining our view that cage-free systems are better for chickens. We want to do more of that. We also just want more time to learn lessons from the grants we've made. We want to build a stronger organization with a more robust operations team and grants management team.

11:57 HK: And finally, there is this major project of cause prioritization. So we picked this initial set of causes that I've listed for you, and I've left some out, but we picked these initial set of causes based on importance, neglectedness and tractability, but before we really bring the budgets up, we want to have a more rigorous sense of exactly what the budget should be for each different kind of giving, and that is definitely a conundrum. We wrote about it at length last year, at the end of last year... I think it might have been early this year, but it's an update on cause prioritization in Open Philanthropy, and there are some really interesting things you get into there, because in an ideal world what you might try to do is you might try to look at each area you're working in, farm animal welfare, criminal justice, pandemic preparedness, and you might try to estimate, "Okay, for each dollar we give, how much good are we doing, how many persons are we helping?"

12:51 HK: And then you might try to allocate the money to maximize the total amount of good you're doing, which means that as you start hitting diminishing returns in one cause you shift money into another cause. That's kind of the first order, a logical way to do it, and you run into these situations where if you value chickens at 1% how you value humans, then you're going to end up with your numbers saying you should basically put all your money into animal stuff, and on the other hand, if you decide that it's like 0.0001% then you should put no money into the animal stuff, and both conclusions seem a little odd to us, given just what we've seen that we were having exciting results in both areas.

13:26 HK: Similarly, when you think about future generations, if you reduce the odds of humanity going extinct by 1%, how much is that worth? Is that worth 1% times the seven or eight billion people who are alive today? Or is that worth 1% times the 20 billion people who'll ever be alive, or maybe the number of people who'll ever be alive is more like 20 billion billion billion billion billion. [laughter] These numbers have very different consequences. And so, one of the things we've struggled with is we have these few key parameters. How do you value future generations? How do you value animals? And they have a tendency to just... If you try to do a quantitative estimate they take over the whole thing and everything you're doing comes down to that one number, and for a variety of reasons, we don't want to be making kind of a wild guess and then putting all the money into one cause. We've written about a whole bunch of reasons which I can go into if someone asks a question about it, about why we want to be in multiple areas, following multiple world views, multiple sets of assumptions, not just have it all in one.

14:27 HK: And so that has led to definitely an interesting set of questions. And so one of our major priorities before we bring the budget more is to get more clarity on that. And there's a couple ways in which we've tackled that problem. So one of them is the bucket approach, so you can say, "Well, I don't know if I should value future generations at a small number or a big number, why don't we just take half the money and have that bucket act as if it's a small number, and half the money act as if it's a big number?" And then you might make adjustments from there where you might say, "Well, actually, if you had two people and they had different opinions on the far future, they might also make some trades with each other, that they might try to make a deal where each one gets more money when they're in a better situation," and things like that. So that's one angle we're taking on the question. And then another angle is just trying to lay out tables to make the trade-offs clear. So one of the things we've been working on recently is an estimate of the value of the last dollar in certain causes. And so, basically what we're thinking is, if we try to spend all of the money on global catastrophic risk prevention, what would be like the worst grant or the grant we made at the end, how good

would that be?

15:42 HK: And then if we tried to spend it on farm animal welfare, how good would that be? And if we tried to spend it on something else, how good would that be? And we try and line-up these estimates next to each other and say, "Well, we have kind of estimated and if we wanted to put maximum money into global catastrophic risk reduction, one thing we think we could spend over \$10 billion on if we wanted, would be research and development into platforms for medical countermeasure development." So in other words, investing in research to develop the ability to quickly develop a countermeasure if there's some horrible pandemic. And we've tried to rough that out in a very rough guessworky way and arrived at a figure of something like, it would cost in some sense like \$245 trillion per extinction event averted or something.

[laughter]

16:29 HK: Which you could think of that as if you spent \$2 billion, it would be some tiny fraction of an extinction event averted and then you can multiply that through, and then we can have a table where we say, "Alright, this is how good we think our last dollar in risk reduction is, this is how good we think our next dollar in animal welfare is. Which one do we like better? How much money do we want to put into each?" And a lot of that is ultimately just going to be Cari and Dustin deciding what they want to do. But I think it's a more... It's going to be a more informed and reflective way of thinking about that question than I would guess most philanthropists are probably doing. So that is a major thing we're working on this year. And then the other major thing we're doing this year, and aside from continuing to give, continuing to keep that machine going, of the \$100 million a year and more, is we're just trying to build the organization.

17:18 HK: And so one major thing is building our operations team. We're very excited to have had our Director of Operations, Beth Jones, join recently and we've also hired Danny and Anya recently, our two new operations staff. And so that's been a major focus of this year is just getting ourselves in a stronger organizational shape. And then we're also on a major, major push to hire more generalist research associates here, which are just the kind of people who can work across causes, help us decide how much money to put into each cause, and then who also we think could have the potential to just become contributors in lots of different ways down the line. They may become future program officers, they might become future people who work on learning and impact, which is not a major department yet, but it will be. It's kind of the practice of looking back on how our grants are going.

18:05 HK: And so, that's been a big endeavor, because we see the research analysts as the future of the organization, and we see them as people we're going to invest a lot in, but years from now, they'll be the people that we were really glad to have invested in. So, we've really just put a lot of work into that. We had over 800 applications for the research analysts' roles and we've been just really diligently trying to create good work tests to filter those people in a fair way and probably hopefully pick 10 or 15 to bring to a three-month trial, from which the strongest people will get the roles.

18:42 HK: So that's been a major thing this year. And those are some of the things that we're working on hardest, so just the organization building and then the cause prioritization framework and estimates of where our money should be. And I think at the point where we're... Where we've really made a ton of progress on those, and we feel like not only organizationally strong, but intellectually strong, I think then we'll look to raise the giving level hopefully, with all that knowledge in hand. So that's the brief update on where we are and then I'm happy to just take questions for the rest of the time. Yup.

19:14 S2: Do you have a way of factoring in something like climate change, things that are related to human population over time? Like the sort of tangible value of the world itself?

19:26 HK: Sure. I might have to ask for a clarification on that question. And this is in addition to climate change, are we valuing things like population? Can you elaborate?

19:34 S2: It sounded like you were coming up with a base that was mostly around human cost, and so, with something like climate change or these global catastrophic events, how do you evaluate the cost of other things besides human lives?

19:48 HK: So, are you saying, when we think about reducing the risk of extinction, are we valuing how extinction would affect things other than humans? Is that kind of what... Yeah. Yeah, that is definitely something we've thought about and definitely something we've talked about. I think one question is, there's a lot of non-humans, what happens to

them in the event of an extinction? Our current take on this is kind of like, if you try to model how good it is to not go extinct, generally, you just, by default, the equation pops out, these just massive numbers that imply that whether you're counting non-humans or not, it just implies that that's kind of the highest leverage thing to work on, and that's the most important thing to work on. And then, there becomes this question like, "Okay, is averting extinction good, is the sign positive?" And we think, yes. [laughter] And that is maybe a longer conversation, but it's definitely something we think, yes.

[laughter]

20:50 HK: So, I think it hasn't been... I don't feel like our calculations are super sensitive to a lot of detail there, it's more like, do you philosophically want to accept this sort of projection that says preventing extinction leads to this huge amount of value, that kind of is probably going to be more than any other numbers you're working with? Or do you feel like there's something wrong with that argument, and you don't want to go all the way in on that and you want to reserve some money more for helping persons or animals today? And that's... And when we try to divide up the money, those are the kind of questions we're asking. Cool. Yup.

21:28 S3: What kind of work are you doing in the scientific research area? And I have a particular question of, whether you're looking into eradicating certain diseases, like whether ones that affect developed world or neglected childhood diseases, maybe trying to eradicate malaria and stuff like that?

21:46 HK: Yup. So, talking more about our scientific work, and one of the questions is whether we're trying to eradicate malaria. So, I'll start off with yes. [chuckle] So an example of a science grant is Target Malaria. Basically, there's a technology, a biological technology, called gene drives that may make it possible to modify mosquitoes, certain strains of mosquitoes, just a subset of all the mosquitoes, so that they either can't transmit malaria or may not be reproductively viable or what not, and have that gene spread rapidly throughout the strain of mosquitoes. And there's a theory that if you could design this right, and release it you might... This might be the way... This might be a way that works to eradicate malaria, which is kind of a very... It would be a pretty epic thing if it happened, in that, to date, I believe that humans have only eradicated two diseases, one of them is an animal disease that only affects animals, the other one is smallpox, which had an effective vaccine. The idea of taking something like malaria that is that widespread and that prevalent. We tried to wipe it out before, and we failed. And there is no very good vaccine. The idea that you might be able to wipe that out is... The toll of that disease is just so enormous. It's disease, it hurts people economically. Just in terms of deaths of children, it's in the hundreds of thousands per year.

23:09 HK: And one way that we thought about this grant is if we could speed up the day that this technology is deployed, every day that we speed it up could be like 1,000 untimely deaths of children averted. And so, it's a pretty exciting goal to have. And what we saw is that there has been a lot of exciting work done on this and it's been supported by the Gates Foundation and all that is great, but we felt... We saw more opportunity to kind of make things move along in a way that is both faster and safer, and is also getting ahead of and being aware of the potential downsides of this, and trying to make sure this thing doesn't get deployed before people have carefully analyzed all the impacts it could have on the ecosystem, but also trying to make sure we get that analysis done, so that if we are able to eradicate malaria, it doesn't have to wait a day longer than necessary. So that's an example of a science grant, the general kind of formula that we use for our science grants.

24:00 HK: We have two science advisors, Chris Somerville and Heather Youngs, and generally, what happens is, we have a list of topics that we think, often on a hunch, sometimes on data, that we think are kind of important and neglected. So just topics that for whatever reason might not get as much attention as cancer. So malaria would be one of them, because it primarily doesn't affect the rich world, it primary affects the developing world. There are other causes, like chronic pain, where there's just an enormous disease burden, but for whatever reason, it just doesn't... It doesn't seem to capture the imagination. You don't have to walk for chronic pain that you might see for cancer or even heart disease. And so, we have a list of areas of interest. And then, Chris and Heather are very broad generalists, and what they do is, they look at an area, they get up to speed, and then they see if anything really pops out, if there's anything that looks unexpectedly great, unexpectedly exciting. And Target Malaria is one example of something that they found. Another example is, they looked around for things that might prevent a pandemic, and they came to the view that we might be able to create a universal antiviral drug from existing drugs, actually from existing cancer drugs, because we might be able to interfere with part of the body's machinery that nearly every or perhaps every virus uses to replicate itself.

25:16 HK: And so, that could be... A universal antiviral might not be as good as it sounds, 'cause it would have side

effects, and you wouldn't want to take it for every time you had anything, but it could be really useful for an unexpected really bad pandemic. And so, occasionally they see something that really stands out, and then we make a grant, and other times they don't, and we're comfortable just saying, "Okay, now we've learned a bit about that area, we'll come back to it if we see something stand out, and now we're moving on to the next area." And I think science has been a different thing from our other program areas, and I think in some ways it takes more specialization and in some ways, it takes less. So, we really need to work with science advisors who have a really deep and long-running scientific knowledge, and know how to navigate that world, but also, when we find an appealing grant, we generally believe that we're able... That the whole system is structured enough and set up well enough that we often believe we're able to commission external reviewers on a reasonable timeline and get a pretty good sense of whether the grant is as good as it seems on a reasonable timeline.

26:15 HK: And so, that's why we're able to be opportunistic in science. It's a style that works for our advisors, and that we think works and in most other causes we think there's just... Any given grant, it's like, the story behind it is so complicated and has so much soft info, and so much stuff that's hard to dig up, that it's really good to specialize and it's really good to have just like... If you're working on farm animal welfare, you want to have at least one person who just lives and breathes farm animal welfare. So we do things a little bit differently there. Yup.

26:42 S4: So you said that some... The advice that Open Phil gives is not, at least not necessarily, followed by the Foundation or even if it generally is. So I'm curious, how often it happens, this divergence, and also, if there's divergence, would Open Phil give the opinion Open Phil had based on research, or not discuss it, or present the Foundation's view on that type of thing?

27:10 HK: Sure. So, the relationship between Open Phil and Cari and Dustin, how does that work, and how much disagreement is there, and how is disagreement handled? I mean, there are... We are formally separate, certainly Open Phil and Good Ventures are formally separate, although Cari and Dustin are also major principals at Open Phil, the organization. They're separate because we want Open Phil to be all about a set of people who wants to find the best giving opportunities, and generally, that's also what Cari and Dustin want to fund, but if at some point we disagreed, or if Cari and Dustin just wanted to do more on things they personally value, we'd have the kind of the separate organizations and could do it in separate worlds. The way it's worked so far is that the vast majority of our recommendations are followed, not every single one, the vast majority of them. We generally loop in Cari and Dustin a lot, we have a lot of conversations with them. It's not like... We're not arm's length, we don't send them something that they've never seen before, we tell them what we're thinking and what we're interested in.

28:11 HK: And there are sometimes grants that Cari and Dustin want to make that are not Open Phil grants. Sometimes we help them with those, because we're partners, but they are not Open Phil grants, and they just go up on the Good Ventures website. An example of this is, Dustin has a personal interest in averting the burden of Alzheimer's disease, and that is something that we helped with, but those grants generally are just going up on the Good Ventures website. And then in terms of if Open Phil makes a recommendation that Cari and Dustin aren't accepting, we continue to believe, and to say when we... When relevant that that's what we believe. But the Open Phil website is not like the GiveWell website, it's not really out there as a resource for people googling around for where to give. Our model of marketing is much more about a small number of contacts with very high net worth folks, so we could definitely take a grant that Cari and Dustin didn't want to fund and shop it around to other funders that we know. But generally, it's not... We're not going to publish a grant recommendation out into the ether for whoever wants it, because our website doesn't have enough of those to make it an interesting source of them. Yup.

29:21 S5: So, it seems to me a part of the research you're doing is like, almost like philosophical questions, so, how useful has the academic philosophical literature been for you?

29:31 HK: Sure. So, a lot of the things you're describing sound like philosophy, how useful has the academic literature been? Super useful. On the relevant questions, so, one of the questions when you're valuing future generations, is you might ask the question, "Well, if we go extinct, how good is that?" It's a question I've asked a couple times tonight. [laughter] One question is... "Well, there's all these people who might or might not get to exist. Is making it the case, those people get to exist, is that like saving lives or not? Is that like something that we should value similarly to preventing an untimely death or is that something that we should just be like, 'I don't care about this?'"

30:07 HK: And I think different people have different intuitions on this and a lot of people approach a question like this by saying, "Well, that's a subjective question. I'm just going to go with what seems right to me." But it's interesting that there's a pretty decent, pretty mature philosophical literature on this. It's a field called Population Ethics. I and others at

Open Phil have read, in some cases written, in my case just read, hundreds of pages on this. One of our employees did a dissertation on a very, very related topic. There's a lot going on and there's a lot of really interesting stuff. And one of the things that's interesting is a lot of people's first reaction is yes or no, but then if you hit them with some thought experiments and you say, "Okay, well, you don't care about helping people exist, but do you care about preventing existence of people who might be suffering horribly?" And then they'll say, "Yes." And then you say, "Well, if you think that it's bad to create someone who suffers horribly, but not good just to create someone who's happy to exist, then does that mean that probabilistically it's just always bad for new people to come to exist?"

31:08 HK: Because it's all downside, no upside. [laughter] And then people will be like, "Huh." [laughter] And actually, actually it turns out that there's... This is pretty wild, but there's these things called impossibility theorems in population ethics, where someone will just say, "Look, obviously, it's bad to have someone come into existence who's suffering. Obviously, it's also bad if this happens. It's good if that happens." It lists six things that all seem obvious. And then, the scholars are like, "And these contradict each other." [chuckle] And so, it turns out there's no population ethics, and this is a generally accepted view, there's no consistent population ethics that doesn't have something wrong with it. And by wrong with it, I mean just something that most people would not feel good about. [laughter] And so, I think that's pretty good, pretty interesting. And so, one of the things I've learned from population ethics from understanding and digesting the literature is that I don't want to totally just trust my first intuition.

32:07 HK: I actually think that it's a really tough question and I want to wrestle with the different ways my intuitions are contradicting each other and try and think about in the end what I want to swallow, and what I'm most comfortable with, and we all do that differently. So, I have my views. The people working on this have their views. Cari and Dustin have their views. But I think that reading that literature, understanding how tough it is, understanding there is no good answer, thinking about which bad answer you're most comfortable with, watching how other people grapple with that too, how you want to deal with their opinions. They might be right; you might be wrong. I think it does leave you in a different place. I think another example of this is like, you might ask, "How much should we care about chickens?" And a lot of people would say, "Well, not at all," and you might say, "Well, why not?" And they'd say, "I don't know, because chickens aren't conscious or chickens don't feel pain." And they'll be like, "Alright, what does that mean? How would we know if chickens feel pain?"

32:57 HK: And so that... I mean, we've got a several hundred page report on this now that incorporates a lot of the literature from philosophy of mind. And the answer is like, again, kind of similar to population ethics, it's sort of like, there's no good answer. We don't know, but I think reflecting on this has really made us move into more of a gray area and feel that there's a great deal of uncertainty, we shouldn't just go with our initial intuition. So, in the end, Cari and Dustin are going to decide what to do with their money, and they're going to do it based on what feels right to them. But I think going through this process of engaging with that literature, recognizing when you're making a judgment call that other people have weighed in on and thought about, I think it's going to lead to more considered decisions and hopefully, ultimately better, better, more reasonable ones. Yep.

33:44 S6: So, I have a related question. So, if figuring out how to deal with moral uncertainty or figuring out which more of these that are more plausible is so important to you, do you work with philosophers or do you fund fundamental philosophy research in figuring out these things or do you try to do your own practical reasoning?

34:02 HK: Sure. So if this stuff is so important to us, do we fund the philosophy or do our own practical reasoning? First off, you mentioned moral uncertainty, and this is kind of interesting that I... I talked about two problems just now that have a good sized philosophical literature. So population ethics, is it good for people to exist? And then philosophy of mind, like what counts as a mind, what's conscious, would a chicken count? There's a third question that's very important to us, which I've kind of hinted at, which is, what happens when I think maybe the best thing to do is to help chickens and maybe it's not, how do I handle that? Like, should I assign a probability over here and a probability over there and multiply them through and take the expected value? Should I do something else? Should I split the money 50-50? That could be seen as a subset of the problem of moral uncertainty. This seems like a philosophy problem. There's really not much that's been written on it, like a few philosophy papers. One of the people whose written the most about it is Will MacAskill, who is one of the founders of the Centre For Effective Altruism, and so, has had a lot of interaction with us.

35:03 HK: And so, that's kind of a younger field. And for stuff like that, I think we do... Our general approach is, we definitely want to read whatever is out there and we want to find anything of quality that might help us, 'cause it's always better to go off of work that's been done than to try to do it again ourselves. When there's a really thin field on something, like moral uncertainty, we may have to come up with ad hoc solutions, and a lot of times those solutions are



not very elegant and aren't very philosophical and are more just us trying to avoid a disaster. So, the kind of method of saying, "Well, half the money is going to act this way, and half the money is going to act that way," I don't know that a lot of philosophers would be very enthusiastic about that. But as a philanthropist who doesn't want to get stuck in one bucket forever before we've unconfused ourselves, I think it's a practical move for us for a whole bunch of reasons that we've written about.

35:50 HK: And then, the final question, so that, yeah, there's doing our own work when we have to. There's reading the literature when we can, and there's funding future work, and we are interested in funding further work. I think a lot of that comes down to what are the opportunities, what is shovel-ready, who is available to be funded and then to the extent that we've run out of things to fund, do we want that to be one of the small number of areas where we really invest our staff time in creating more things to fund? And so far, the answer has been, we fund things that are available to fund but we're using our scarce staff time on other things, on other causes where we feel that there's not enough attention going in. So that's so far, but one of our major grants is... And I don't know if it's totally finalized yet, but it's a grant to the Global Priorities Institute and they are working on questions just like this, they kind of grew out of some of the Effective Altruism stuff, and we're funding them. Yep.

36:44 S7: Can you talk more about the medical countermeasures, like why it's a good baseline, and whether it considers future lives?

36:53 HK: Sure. So, why is medical countermeasure a good baseline and does it consider future lives? Why it's a good baseline, we were trying, again, to think... One question you can ask is, "Should we make this grant or not?" And we often think that question is... Really, the right way to ask that question is to say, "Should we make this grant or should we save the money and give it at the end?" Sort of if we don't make the grant then the money effectively funds the last thing we would have funded otherwise. And so that's the last dollar question. So we're always very interested in having a model of like, if we don't make this grant, then what is the money going to do? And that reduces to the question of like, "What's the last dollar we're going to spend... How good is that going to be?" To answer that question, a shortcut we tried to take was thinking of things we could do that would just cost a ton of money and might be pretty good.

37:40 HK: And so, obviously we could try to list the next three thousand things we're going to fund, but that would be very hard, but a thing we can do is we can say, "If we can come up with something that would be pretty good and would cost a huge, huge amount of money, then at least we know we can always fund that with the last dollar." And so then at least we know that if a grant is worse than that, we probably shouldn't make it and if a grant is better than that, we plausibly should make it. So we made a whole list of things where we thought maybe we can spend a ton of money on this, maybe we can spend a ton of money on that. And the medical countermeasures was just an example of something that looked very expensive and pretty good on kind of a long term as a metric. And we tried a bunch of stuff and that was one of them, not the only one.

38:24 HK: Does it consider future generations? Like I said, we evaluate these things in terms of at a first approximation, the cost per extinction averted, and then you can break that down in other metrics. And so, \$245 trillion per extinction averted implies one thing about the cost per life saved if you think that's worth eight billion lives, and another thing if you think that's worth a billion billion lives and we actually have a spreadsheet where we have both, and so we say, this kind of work is worth this much if you take this philosophical view, and that much if you take that philosophical view, and this other kind of work, which is maybe like farm animal welfare, is worth this much, if you take more of a near termist animal-inclusive view and worth that much if you don't care about animals and worth that much if you only care about the long-term future, and then we just look at the table. And at least then we've got a lot of stuff on one page to think about and think about how much money Cari and Dustin are going to want to put into each kind of thing. Yep.

39:20 S8: What is your relationship with charities after you grant them money, do you have seat on the board or how do you make sure they're [39:27] \_\_\_\_?

39:29 HK: Yeah, what's our relationship with the organizations we make grants to? It varies a lot. The basic structure of Open Phil is we pick the causes, then we pick the people, the program officers, and then the program officers really have a ton of autonomy, and they kind of run their own portfolio within their own cause, and so by default it's the program officer's job to keep up relations with their grantees, to check in on them. We try to give them guidance, general guidance, lessons we've learned from all across, so we generally encourage them to spend a ton of their time just getting to know people and having good relationships and having the kind of relationships where people trust them. That's a major emphasis at Open Phil. In terms of checking up on grants, we have a schedule of every time you make a

grant, you say how often you're going to check in on it. Generally, my view is that as a funder, it is grantees care a lot what our program officers think of them, and if anything, too much, if anything, it can be difficult to get honest feedback from them, because they don't want to be on the wrong side of a funder.

40:35 HK: It can be difficult to give them feedback because they might just work too hard to act on the feedback instead of pushing back. So, in my perfect world, we would say, "Hey, your organization seems like it should be doing A instead of B," and then they would say, "No, we think we're doing the right thing, but thanks for the input." But what happens in the real world, is a lot of times they say, secretly they're thinking, that's horrible advice, and then they say, "Oh, that's great, we'll get right on that 'cause we want to be..." So I think a lot of the challenge we have is trying to have healthy two-way feedback with our grantees, and not walk too heavily. I don't think we generally need board seats in order to enforce accountability. I think in many ways we're worried about too much accountability in some sense. And just them wanting to get their grants renewed is plenty of accountability. So board seats are fairly rare. We've done it in a couple cases for very specific strategic tactical reasons, that we wanted a board seat because we wanted to make sure we had a say in something particular. But in general it's the relationship with the program officer and the prospect of future money is plenty. I'll take one or two more questions. Yeah, in the back.

41:49 S9: Do you ever encounter a funding gap where there's a potential for profit and what do you do in that scenario, so...

41:57 HK: Sure, do we ever encounter a funding gap where there's a potential for profit and what do we do? This is sometimes called like mission aligned investing or impact investing. The way we think about this is, A, we're not an impact investing shop, we're also not anti-impact investing, it's not really the main line we carve things up along. So basically, if Lewis is working on animal welfare, his job is to find the best animal welfare stuff to do. If that happens to be grants, we make grants; if it happens to be investments, we make investments. So that's the theory. But the other thing I'd say about our attitude toward investments is we consider our specialization to be doing good and finding ways to do a lot of good with money. We think our specialization is not making money, and so if there's an opportunity that if it works, it'll make a ton of money, our default is kind of like, "Well, lots of groups are interested in making money, and so that's unlikely to be neglected for the kind of reasons we would care about." [laughter] If no one is investing in this it may actually just not be as good as it looks.

43:02 HK: Whereas when we look at a charitable opportunity, we say, "Boy, this would help a lot of people, but no one's funding it," that's not very surprising to us and a lot of times we go ahead and fund it. And so in general, we have a default that we're not out there looking for the investment opportunity. Occasionally, we'll see it. And the way we think about it when we see it, is we usually assume that any investment opportunity we're interested in, we would assume would be worse than a normal investing opportunity. That's kind of the whole point. And so the way that we would think about it is we would think, "Okay, we're going to put in X dollars and in expectation we expect to recover some fraction of that." We don't expect to make more, because that would kind of imply we were beating the market in a way we haven't set ourselves up to do. In expectation we'll get some of it back, and the difference between what we put in and what we expect to make back is the effective size of the grant.

43:50 HK: And so, if we expect to lose half our money, if we're going to put in \$10 million somewhere and expect, on average, to lose half of it, we treat that as like a \$5 million grant, and then think about how much good we're doing and compare it to everything else we're doing. We have done this at times. So we have an investment in Impossible Foods, which makes these burgers out of plants, that supposedly taste like meat, I think they're pretty good at it [laughter], but I'm very conflicted, so don't listen to me. And yeah, we have done impact investment but we're usually through the lens of... We invested in Impossible Foods because we wanted... We wanted them to do more work on meat alternatives and we were happy with what our money was paying for, even if we never saw most of that money again. Take one more... Yeah.

44:36 S10: If you could add like one more thing to importance, neglectedness and tractability, what would it be?

44:43 HK: Sure. If I could add one more thing to importance, neglectedness, tractability. First comment is those are just heuristics. So, those are like the ideal thing to do would be to... Kind of like how many persons are we helping per dollar and just pick everything that way. Importance, neglectedness, tractability are just kind of like practical tools for honing in on good causes, when you're very uncertain and you lack a lot of information. I think something that's come up recently as a candidate to be added to that list, so I don't know this is my first choice but it's something that popped up, 'cause it's come up, is feedback loops. So is this a cause where we're going to be able to learn as we go, is this a cause where we expect to see meaningful results and learn things on a reasonable time frame?

45:23 HK: So I think all else equal I'd rather, I think there are advantages to farm animal welfare and criminal justice reform, where we can expect to see things and we can see if we're doing well or poorly, and we can see the return on what we're doing as we go, and that does have some appeal over a cause like pandemic preparedness, where it's like every year that goes by, we're like, "Yep, we didn't know, we called it a pandemic again." We don't really know how good our giving is still. And so I'm glad we're working in those causes that are incredibly hard to measure, because I think their importance, neglectedness and tractability is amazing and makes up for it. Or at least the importance and neglectedness does. But I think I think the feedback loops would be an interesting thing to add there, and I certainly think they matter. I wouldn't only want to do causes that are strong on that, but I do think they matter. Cool, well, that is it, so thanks for your questions and I'll stick around for 10 minutes or so. Thanks, everyone.