GiveWell San Francisco Research Event December 4, 2018 – Open Philanthropy Project

SF Research Event 2018-12-4 Open Philanthropy

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: Record on these things. So I think you guys all know the drill of the recording. But yeah, it's being recorded. So, and I am going to repeat questions for the recorder. So yeah, I'm going to about the Open Philanthropy Project. Obviously, it used to be GiveWell but I'm not there anymore, so that's what I'm going to focus on. I'll start just with the basics of where we came from, what we are, and then I'll give an update on what we've been doing over the last six months or so and then I'll take questions.

00:27 HK: So, first off, just what is the Open Philanthropy Project? I think a lot of you know this, but I'll go over it pretty briefly. Co-founded GiveWell with Elie and the idea was, find the best things to give money to, but it was very targeted toward people like me and Elie who are individual donors, kind of short on time, looking for somewhere we can write a check. And after about five years at GiveWell, we ran into Cari Tuna and Dustin Moskovitz who were... Dustin being one of the co-founders of Facebook, trying to give away a multi-billion dollar personal fortune within their lifetime and just felt that they were facing a very similar question to GiveWell in some ways, which is: How do we give away a bunch of money to do as much good as possible? And very different in other ways, just looking at their situation just felt like it called for more of an approach that could involve, for example, getting into areas where you might be creating new organizations or you might be making grants that really transform what organizations do and where you could have the ability to hire your own experts, get deep into causes and do things that would be very hard to explain on a website and a very poor fit for a lot of the people that use GiveWell.

01:35 HK: And so, in some ways, GiveWell and Open Phil sit at sort of opposite ends of a spectrum and I personally like both of those into the spectrum better than I like the middle. So, I still donate to GiveWell's recommended charities. So, I think GiveWell is, among other things, but I think GiveWell is really giving you the recommendations that I think you can make the best case for, that one can read about and understand all the reasoning that goes into them and feel that all the homework has been done, and you're as close as you're going to get to knowing what you're getting with your dollar and knowing that it's the best deal for your money. And I think that is one model that obviously I was like, "Set it up about to co-found," and I think it's great.

02:19 HK: Open Phil is more of a hits-based giving philosophy, so... And we have a blog post about this called hits-based giving. But the idea is that it's a little bit more venture capital oriented in a sense, where the idea is that you might make a large number of grants and the vast majority of them might completely fail and then a small number of them might succeed in a way that's big enough to make up for everything else. So we've been very inspired by some of the massive hits in the past for philanthropy, you could arguably credit philanthropy for the Green Revolution that brought like a billion people out of poverty. You could fairly credit it for at least a major acceleration in the development of the pill, the most common oral contraceptive, at a time when no one else was really funding or interested in funding that kind of work. And so, if you could fund one thing that was like that, you could waste an awful lot of money in other stuff and still come out ahead, and still come out looking pretty good.

03:13 HK: So, the philosophy of Open Phil the way in which it's kind of opposite is that a lot of times we're trying to orient everything around getting those massive hits and a lot of times what that means is trying to get ourselves in a position where we're actually making the grants that are hardest to explain. I think grants that are hardest to make a case for and we really set up the organization so that there are people who just... They work in a field like criminal justice reform or farm animal welfare, and they just live and breathe that field, they know everything about it, they know everyone in it, and then they see things that are really great in ways that would just be very hard to justify in a traditional framework and those things get funded oftentimes even when I don't understand them or when the funders don't understand them fully. So Open Phil is very much designed to do things that way and that has led to a major divergence in the style of the two organizations even though the values are extremely aligned.

04:03 HK: So, Open Phil is way less about sharing the case for everything online. We still do try to share more information than other foundations to help people learn from what we're learning, but a lot of the stuff we do has kind of policy aspects and other reasons that we need to be very careful with the information, and so we are considerably more button-down in how much we share. And that does affect the way that I talk about it, so apologies for that in advance.

04:28 HK: In terms of what we've done so far and where we stand, we've generally taken the attitude that we want to... The basic Open Phil formula is step one, pick focus areas. So causes to work in. And we try to pick causes or problems

to work on that are important, neglected, and tractable. So important means that, if you could get a win, it would be a huge deal, it would affect a lot of people, it would do a lot of good. Neglected means there's not a million other philanthropists working on the same thing and tractable means we see some kind of path to victory. And through the process of identifying the most important, neglected, tractable causes, we generated a list of causes and then for each of them, we tried to build a team around the cause. So, when we were decided we were interested in criminal justice reform, farm animal welfare, we have a person in each of those who leads our work. It's Chloe Cockburn, for criminal justice, Lewis Bollard, for farm animal welfare. These are examples and they really are people who just kind of spent their life on that cause, know everything about it, and then we set up our grant making process.

05:35 HK: One illustration of the way we think about it is our grant-making process has this sort of 50-40-10 rule. So what that means is that someone like Chloe she's expected that 50% of her grants, by dollars, would be endorsed in the sense that the decision makers, myself and Cari, are truly convinced by the case for the grant and could truly, this is a grant that I feel like I can explain and defend to someone else. However, another half of her budget can go to things that are merely approved in the sense that I can see how someone would think the grant was a good idea, if I knew a bunch of stuff that I don't currently know and I'm kind of deferring to her expert judgment. And then 10% of her budget is called discretionary and that means there's basically very few controls on it. We stop things if we see red flags.

06:22 HK: But other than that, it's often like a three-paragraph email, if I don't see any red flags, it just goes through in 24 hours and it doesn't matter if I really agree with the grant or not. And that's trying to strike a balance where you have people who really know their area and we think we would slow ourselves down, and miss a lot of great opportunities even just requiring things to make sense to hold in. And I think if we broaden that out and require things to make sense to the whole world, I think we just lose a ton of our best shots at impact. And at the same time, we don't want to go so extreme that we have no idea what our program officers are doing. So that's where the Endorse requirement comes from that at least half the portfolio is going to be stuff that I personally understand, can explain, and can defend. And so that's the basic formula.

07:05 HK: I still think of us as somewhat start-upy, in the sense that our annual giving is below where it will eventually need to be to reach Cari and Dustin's goal of giving away their fortune within their lifetimes, and I think we still have a lot to figure out as an organization before we bring giving up to that level. So right now we give away over $100 million a year but I think it will need to go up further because of the fortune that Cari and Dustin have. And right now, I think we're still figuring a lot of stuff out, which I will get into. But I would say, we're kind of... It feels like a bit of a mid-stage organization. We have been going since 2013 and I think we've probably got several years to go before we're going to be ready to kind of get to maturity, get to the maximum giving we're going to do and feel that our frameworks are all set. So just updates on some of the grants we've been doing and some of the work we've been doing, and some of our causes, and then I'll talk about some of the cross-organizational issues that we're dealing with and how we're trying to develop.

08:07 HK: So in criminal justice, one of our major themes has been prosecutorial accountability. So we basically feel that the US incarcerates people at a rate that is dramatically higher than other rich countries and is also dramatically higher than is reasonable by more or less any standard that we can come up with. David Roodman who also has done work for GiveWell did an extremely thorough review of the evidence, he replicated and kind of refuted a bunch of studies and ended up concluding that on the margin, there is, his estimate of the reduced incarceration, his estimate of the effect on crime is just zero. So it's just kind of like a free win. That's his best guess. It could go different ways, but that's kind of the model we're working with. So reducing incarceration on the margin, we generally think of is a good thing. And in general, we have this view that currently what a lot of the influence in criminal justice comes from prosecutors. So they're the people who are in charge of whom to charge, how aggressively to charge them, and the traditional way that prosecutors tend to be judged by the voters because the head prosecutor's often elected position. Traditional way that people tend to think about a good prosecutor is someone who is tough and someone who gets a lot of people locked up and wins a lot of cases.

09:28 HK: And we're interested in the idea of people thinking of a prosecutor as something different as someone who balances cost and benefits and tries to do the best thing for their community, which sometimes means being less harsh, locking fewer people up. Great example of this is Larry Krasner who's fairly recently elected prosecutor in Philly and has really been revolutionizing things. Just putting out memos saying like, "For certain crimes, if you want to assign a sentence above a certain recommendation, you've got to explain why the stage that spent $50,000 a year for this person to be in prison and why that is a good deal for our society," and a whole bunch of other stuff like that. So they have a lot of power, a lot of discretion. And one of our interests has been encouraging just a culture of thinking about prosecution differently, and we're excited because in the recent elections, there have been a lot of wins for more kind of progressive

in that sense, prosecutors. And we think that a lot of the groups we fund are therefore going to have more sympathetic people listening to them. Another major group we support is the Alliance for Safety and Justice, we've been supporting them for a long time, they've gotten some really impressive wins on sentencing reform, so just converting various crimes to have lower penalties.

10:39 HK: They did have a really exciting measure on the ballot in Ohio that lost pretty badly and that was disappointing. And that's going to be one of the things that we're looking at over the next couple of months, to just be like, "What happened there and is there anything we can learn from that?" So that's criminal justice.

10:54 HK: Farm animal welfare is a cause where we believe that animals are just treated incredibly horrifically on factory farms. I'm not personally an animal lover, but I've gotten more interested in this issue, as I've seen just how bad it is. And when we got into this cause with Lewis leading the way, we funded a bunch of groups that were campaigning against corporations and kind of saying, advocating for them to at least declare that they'll go cage-free with their egg-laying systems, which is, it's still pretty inhumane, but it's a welfare improvement from the current battery cages that chickens are kept in. And we came in when there's already momentum, there are already some wins, so I think we've only been part of this story. But at this point, there have been pledges for cage-free that have kind of taken over all the major grocers and fast-food companies in the country and when we look at our... It's kind of interesting to do a comparison because I'm not a big fan of chickens inherently or emotionally, but you know they give off figures of a few thousand dollars per human life saved and when we're estimating our impact on chickens, it's more like 100 chickens spared from being in a cage per $1.

12:08 HK: And so, even if you try to kind of adjust that in various ways and consider that it's not the same as being dead and do the various suggestions you could do to compare to a human, it still comes out looking like a very, extremely cheap way to help beings if you consider chickens relevant beings, and that's one of the things that we struggle with. I think it's an open question and we do some work there because I think if the philosophy works out a certain way and if you care about chickens a certain amount, I think that some of the best philanthropy around and one of the most amazing deals that if it doesn't, then I guess we're just not doing much there. And then some of our more recent work, we've been making a lot of grants recently in East Asia, that's an area where we think it's much further away from any kind of tangible victory, but also the stakes are enormously high because a lot of the meat consumption is going to come from East Asia in the future. And then we also... We were recommending gifts to the ballot measure campaign. There's a recent California ballot initiative that we thought it was extremely important where now you want to sell eggs in California, they have to be cage-free. We see that as a very important piece of keeping companies to these pledges and making that dream of a cage-free America a reality.

13:23 HK: So, that's some of the recent stuff. We have a scientific research program, so the theory there is just a lot of philanthropies' biggest wins in the past have been by funding breakthrough science that maybe the government won't fund. We have Chris Somerville and Heather Youngs leading that work some of the recent stuff we've been putting work into. We've made a major investment in a company that is working on cancer diagnostics and we're trying to get them to also work on virus diagnostics because of our interest in biosecurity and pandemic preparedness. So, we believe that another way to help a lot of people per dollar is to prevent the global catastrophic risks that could really affect not only the whole world today, but a whole large number of future generations. And that's something where, depending how you do the numbers, they could work out kind of absurdly good.

14:12 HK: And so, building more tools to prevent a worse case pandemic is a major interest of ours. And speaking of biosecurity we haven't been as active on that recently because we've had a change in leadership. So basically our program officer left a couple of months ago after being with us for a couple of years making some grants to strong institutions like Center for Health Security. We made an offer to another one, it was accepted, but that person is not starting until April. And so, we're kind of in a holding pattern on biosecurity and pandemic preparedness right now. And then the final one I'll put it in a few words about, is potential risk of advanced AI. So this cause is kind of a can of worms to get into with the basic idea is that we are very interested in any kind of work where we could imagine ourselves preventing a disaster that can affect many future generations and that is kind of a matter of numbers. Just like if you want to help the most people per dollar, anything you do that can ripple through the whole future is a big deal. And if there's anything that could ripple through the whole future one of my top guesses, although I will talk about some other ones, is AI.

15:20 HK: If we got to the point where the kind of work that human scientists do, making technological breakthroughs, could be automated itself and so you could dramatically speed up scientific and technological development, that could just be a massive change for better or for worse. We believe that could happen sooner and more quickly than a lot of

people are imagining and that it could go well or poorly, it could lead to a massive concentration of power, it could lead to a whole lot of other problems that are, when we make a list of things that can ripple throughout the generations, that one ranks right up there. And so, a lot of what we do in AI, is we try to support the growth of fields, we try to support an academic field around AI safety research, building systems that behave reliably, understanding not just how to make an AI system sort of smart and capable, but how to make it understandable and well-behaved and act as people are expecting it to act consistently and act as people want it to act which is not exactly the same as having something that's very good at winning games. And so, that is something that we just have a general interest in.

16:27 HK: Some of the recent work on there, we have a fellows program for early career AI scientists where we basically provide a lot of support and then we ask that they be interested in and consider safety and we're trying to make that a more... A kind of an easier field to enter as a young scientist, an easier thing to do is to say that you want to work on these safety and reliability issues instead of just capabilities. And one of our major recent projects has been organizing... Not only renewing our fellowship and picking a new class of fellows, but organizing a retreat for our current class of fellows to get everyone together and talk about what are some of the potential social implications of advanced AI, what are some of the potential risks and what can they do about it? So that's some of our recent work.

17:11 HK: In terms of the overall themes of where we're trying to take the organization, the priority all year pretty much has been hiring. So in terms of our evolution, we, at our first couple of years, we were just picking causes to work on. We had a whole year then of hiring, just getting a few program officers and then we spent a couple of years scaling up grant making and getting into our kind of $100 million a year level. But now we have a whole lot of other needs. So we hired a director of operations this year, we're building out our operations team just to be able to support a larger organization, and then a couple other things that we need to get better at and that are becoming very important for us.

17:49 HK: So, one of them is cause prioritization. So, we already picked causes to enter that looked interesting, that looked important, neglected and tractable, but before we ramp up to maximum giving, we're going to have to decide how much money goes into each cause each year. And I think that is a much tougher decision that requires us to look at a lot more detail. And the way the GiveWell does it, GiveWell looks at all these charities and they say, "How many units of good can you get?" And they measure that different ways per dollar spent. We need to ask the same question and it's an even more daunting question for us because some of the units of good are in years of incarceration reduced, some of them are in chickens not in cages, some of them are in slightly reduced probabilities of some horrible catastrophe that affects God knows how many future generations of people and how many people in them and how many animals too.

18:42 HK: So comparing all those things is a huge challenge, and we do have a tentative research agenda, and we do have a tentative take, but a lot of what we need to do is just hire more people and get better at that, and have more of a fully fleshed out framework the way that GiveWell does, of saying, "This is how much good we think we can accomplish in each cause and that's why we're putting a certain amount of money into each cause." And then another function that we're building is self-evaluation. So we've been giving it scale for a couple of years now, and I think we need to start looking at what we think, how our wins have been, and really taking a critical look at them and saying, "How much impact have we had? Is that in line with how much impact we thought we would have? And what does that tell us about how to adjust our estimates of how much good we can do in different causes?" So the key to building out those two functions, I think they're both really poorly defined functions. I think you can't really find anything at other foundations that looks like what we're trying to do there.

19:34 HK: Other foundations do self-evaluation, but I think they do it very differently from us. We're very interested in getting that quantification of how much good did we do and what were the really big wins? And so we think we need generalists, and so we spent a major project for just a lot of people at the org all year, has been trying to hire generalists, and the way we did it is we did a big call for applications, we got over 800 of them, and we ultimately invited about 15 people to a three-month trial that just wrapped up on Friday, and so this week we have a whole bunch of decisions to make about who gets offers. And so that's been a major preoccupation this year. So in terms of the evolution of the organization, we've reached the point where we're giving at some scale, but I think we need to get better at nailing down what are we actually accomplishing, and what I think we can actually accomplish going forward, before we have the kind of principled approach that makes us ready to bring giving all the way up as high as it's ever going to go. And so that's basically where we're sitting in our evolution. So that's where we've come so far, and yeah, it's been about 20 minutes, I'll just take questions the rest of the time.

20:38 S?: So GiveWell has sort of taught me to be skeptical of US charities. So I'm a little surprised that even policy levers, GiveWell welfare and other areas, criminal justice that... I guess I expect more cost-effective options if you're

looking outside the US [20:57] _____.

21:00 HK: Yeah, so GiveWell has taught you to be skeptical of US charities and why are we doing so much stuff in the US. So I think that's a great question. I would lay out the framework, sort of like, you're a philanthropist, you're trying to give away money, and you want a big return, you want to help a lot of people per dollar spent. What are some ways you can do that? And I would say the main kind of concepts that we would bet on the most as ways to get an outsized impact, one of them is sending the money to poorer places where it can go further, and I think that is the main way that I would say GiveWell gets extra impact out of its dollars. And so, to give an example, if you... The simplest version of this is, is if you give to GiveDirectly, then you're probably giving to someone who's roughly 1% as high income as the average American, and so your money is going something like 100x as far, as if you just kind of kept it in the US and gave it to random people, and that's good.

21:55 HK: Another way you can get a multiplier is I think policy, where if you can spend... You spend some money, and then you can cause... You can increase the odds of a change in policy that then is from this huge player, like a government, or actually in the farm animal case, sometimes it's a corporation. But someone who has far more money than you do and makes far more consequential decisions, that's another kind of multiplier. And I think when we've done the estimates for the US, I think it's not unrealistic to imagine that you can get 100x return there as well, or even 1000x on your dollars, which I think somewhat corresponds to the range of GiveWell where 100x is GiveDirectly, and 1000x might be the charities that GiveWell's estimating to be 10x as good as they are. So that's a multiplier. There's some other ones too. I think when you're doing research and you develop new ideas, those are public goods that anyone can use and the impact is kind of unbounded, and also things that reduce global catastrophic risks, I think you also have this unbounded impact because of the way they can affect the future.

22:53 HK: But those are the major multipliers. And so what I would say is GiveWell is getting this one multiplier from going overseas, we're getting another multiplier from going with policy, and GiveWell until recently has not done as much with policy advocacy, and I think both we and GiveWell are looking at, "Could you combine the two? Should you be doing policy overseas?" And I think we're just coming at it from different directions, because when we first got into policy, that just felt like too much for us to bite off at once, and we wanted to develop an understanding of policy change in a context that we knew.

23:28 HK: And also, I think there are arguments that it's not necessarily going to be better when you go overseas. The US is this disproportionately powerful country whose policy ripples out throughout the whole world, and some of the causes that we work on, like immigration, macroeconomic stabilization, they really are global just by work in the US. But we are thinking if we did more policy work, we would seriously consider going overseas, we just weren't ready to start there. And I think we have seen returns that are at least not incredibly far off from GiveWell. So the distinction I would draw here I would stand by GiveWell if they're saying it's better to give direct aid overseas than direct aid in the US, like homeless shelter versus bed nets, go with the bed nets; I would fully agree with that. And then I think you're going away from apples to apples a little bit when you're starting to compare direct aid to policy. Yep.

24:16 S?: So while we're on the subject of working within the US, you mentioned that East Asia is a more and more consequential place for farm animal welfare. So could you go into detail about what things you're looking at there and whether it's a similar [24:32] _____ like corporate campaigns, outreach, etcetera?

24:36 HK: Yeah, so what are we doing in East Asia? The thing you should know is that when we came in on the US cage-free campaigns, in some ways it was really exciting because we started seeing wins right away, but we were also really conscious that we were riding away from someone else. McDonald's had already pledged to go cage-free, there was momentum. A lot of the hardest part had been done, and we kind of came in and poured gasoline on it. We think we sped it up a lot and may have made a major impact, but it was definitely... We were only a little piece of that picture, in a sense. And what we're trying to do in East Asia is the opposite of that, is kind of be that first player. We're the only farm animal welfare funder in China. And what that means is I think we have to dig in and just be ready to just stick it out for a long time, and it may be 10 years, it might be 20 years before we see our first win, and then see the cascade.

25:26 HK: So, the main animating idea in East Asia has been, we just fund people who care about farm animal welfare, and if they're doing something plausible where they have some kind of path or some kind of impact or building the movement, that's often good enough for us. So we're not looking for the giant, amazing path to victory right now, we're trying to build the movement in the field. There's been all kinds of stuff we've done. We've just... Like anti-leather campaigns and just advocating vegetarianism which isn't something we do as much of in the US, because we don't think it's as promising and effective as the corporate campaigns, but we think it's, for many reasons, could be more promising

in China and India. So, sometimes we're just funding people who care about farm animal welfare, we just get together and brainstorm. So it's been just a lot of activities. And we have been doing this for a while. The most recent set of grants has actually been even more all over the place where it's like trying to find re-granters to just make lots and lots of tiny little seed grants to people to try things. So there's not one major theme, it's more like who's interested in this issue and how can we support them to start building allies and figuring out what works?

26:35 S?: If I can ask a quick follow-up?

26:36 HK: Yeah.

26:37 S?: So I assume that your vision for what that turns into East Asia is kind of similar to what has happened in the US at the post [26:45] ____. So, could you maybe give us a little history lesson on how that went?

26:50 HK: History on the US, yeah.

26:52 S?: What was happening before you came into the States, and why did it build [26:57] ____?

26:58 HK: Well, I definitely... I was drawing a pretty broad analogy, what we're trying to do in East Asia and the US, in the sense of we're willing to be patient and be slow and take a while to get the first win. I don't think it's going to look exactly the same. I think in some ways, I'm hoping it goes much better because those countries are still... Their economies are still growing so fast, and a lot of it is like, if you could just get vegetarianism to be a little bit more popular sometime in the next 10 years, you might just prevent a whole lot of meat consumption than in the US, it just kinda feels like it's very hard to turn around that train and the corporate campaigns are the best thing we see. So, but in terms of the history in the US, I don't know, I don't think it's so much that it took a long time. It's not like people were trying to do cage-free campaigns forever. It's more like people have been trying different things for a long time. So, people have been trying to advocate for the rights of animals and convince people to go vegan, convince people to go vegetarian, pass laws, like people have experimented with a lot of stuff, and right around the time that one approach really started winning and catching fire was around when we hired Lewis and then just funded it to go much faster. Does that make sense? Cool. Yeah, in the back.

28:11 S?: Thinking on the 20, 30-year time horizon, how much are you affecting early education so that 15 years from now you'll be able to affect policy?

28:21 HK: Yeah. So, for the long time horizon, how much are we affecting early education so years from now we'll be able to affect policy? We don't... So, education itself is not a cause for us, and that's partly because we just think it's very crowded. So generally, it's one of the most popular causes for philanthropists is to do education, and we don't know that we would have a lot to add there. But I think on the issue-specific stuff, getting the message out there to people who are just in a stage where they might be more open to hearing things and plan the rest of their life around them, that is definitely a thing we're doing. So, for example, with the AI Fellows Program, that is people who are... They're going to finish their PhD or they're going to do their post-doc on our funds, and that can potentially shape the rest of their career. And yeah, a lot of the work that we are doing overseas with animal welfare is about getting messages out there to people. And some of it... I think we are doing some work that's on campuses, campus organizing. So it's a tool that we treat kind of pragmatically, it's like we've got a cause, and we do have a long time horizon, so what's the best way to do it, and it's just different for every cause.

29:26 HK: This might also be a good time for me to just go to one of the questions that was submitted on the forum about climate change. And I think it was pointed out that climate change is one of the most clearly global catastrophic risks that is out there, and why do effective altruists, people in the effective altruist community who talk about doing as much good as possible, why do they tend to not seem as interested in climate change? So here, there's kind of a similar answer to education, which is that we feel that climate change is less neglected than the global catastrophic risks that we work on, generally. It is a much more popular topic with philanthropists. There's a lot more philanthropic money in it than there is in like AI risk or in pandemic prevention and biosecurity, especially biosecurity, so especially anything to do with preventing engineered pandemics; I think there's essentially no philanthropy. But we are... That is one of the major reasons that climate has not been one of the causes that we've gotten all-in on, in the way that we have biosecurity and pandemic preparedness. And at the same time, we do a serious amount of funding in climate change, and so we have spent millions of dollars on that cause. We have treated it as more...

30:38 HK: We have some people on staff who might work on several causes and just look for the most outstanding,

most neglected stuff, and a lot of our work on climate change has been on aspects of it that we think are highly neglected, so especially solar radiation management, just like scoping out, what are the world's options if we're in the worst, worst case for climate change, what are the world's options for reflecting something like that to cool the earth, how dangerous is that, could that be even more dangerous than climate change, what's the safest way to do it, what's the best way to convince people we shouldn't do it if we shouldn't do it? All that stuff. So, I think that's... If the world ends from climate change, I actually think most of the action is in the solar radiation management. It's like, either that the cure is going to be worse than the disease and the world is going to end because of geo-engineering, or that's going to be the key to how we're okay. And so, that's been where our climate change work is focused, but we put a serious amount of money into that. Yep.

31:36 S?: So you mentioned earlier in your presentation that historically, from the history of philanthropy, scientific research has been one of the most promising endeavors, and I was curious, what has been Open Philanthropy Project's process for directing funding to scientific research? Are you guys putting out calls for proposals or...

31:57 HK: Yeah.

31:58 S?: How do you identify things to work on?

32:00 HK: Yeah, so what's our process been on scientific research? We've tried a lot of different things over a bunch of years, and we ended up... We've kind of experimented with two approaches to Open Phil, and I don't think we've really come down believing that one of them is dominant yet. So, we have a blog post that is from a long time ago that's called "Expert Versus Broad Philanthropy". I think that's probably not the right term; I would call it more like "Deep Versus Broad Philanthropy" today, because we're trying to make everything expert in a sense. But a lot of what I've discussed is having one cause, having a person who's really obsessed with the cause, knows everyone, has a budget for that cause, and that's what we do. There has been a bunch of work that looks more like, "Here are five things we're interested in, or here are 10 things we're interested in," and we're only going to make a grant when we see something that's just amazing. Just like we know what the most neglected parts are, we know what the most promising parts are, we know enough to know that, and then when we see someone who's got a great plan and he's a great leader, we're going to fund them. And that's been the direction we've more ended up going with science, where we just felt like we never really were able to get satisfying importance, neglectedness, tractability ratings on scientific sub-fields, so we were never really able to say like, what disease is the most important, neglected and tractable.

33:16 HK: Especially because the tractability is just like... That's been the real stymieing part, is it's like, we found that we're occasionally able... Our scientific advisors are occasionally able to spot something that just looks like really, really tractable, and it's being ignored for some really silly reason. And the rest of the time it's just like an incredible endeavor to figure out how tractable something is, and there could be a lot of variation that we're just missing, because we aren't experts in the cause. And so, we experimented with trying to do cause reports and do sub-cause reports, and pick areas to focus on, but either because that isn't the best way or just because it isn't the best way for our team and our people, what we've instead found is that we look into a cause for a little bit, and a lot of times we'll find one thing that just looks amazing, and that isn't being funded for systematic bureaucratic reasons; a lot of the science funding is government, and then we'll fund it and we'll move on, and then we'll do it again. And so, that's been the basic philosophy that we've used, and some of the things we've done, I mean, we funded work to speed up gene drives to eradicate malaria. So this is a way of genetically modifying mosquitoes so that they preferentially pass certain genes, like for malaria resistance, or for not being viable on... And so you could...

34:33 HK: In theory, if we can get this ready in time, that could be used to wipe out entire strains of mosquitoes; not all of the mosquitoes, don't worry, we'll keep the mosquitoes around, I know people love those, [laughter] but wipe out enough that you can actually eradicate malaria. And that was a case for us, like we looked at it and... I mean, it was obviously an amazing social impact, but it's mostly in the developing world. So, a lot of the support you might normally see from the US wasn't there. Normally, the funder that's most interested in that kind of stuff is the Gates Foundation. They were funding it, but they, for reasons that I still don't fully understand, felt that they had only funded it some fraction of as much as they could to speed it up as much as they could. And so, that was the case where we were willing to put in $20 million just in the hopes that we could help it happen like a year faster or something. And if you could do that, that would be a lot of lives saved, just making that happen a year faster it's going to happen. So that's an example of something we did.

35:27 HK: I gave the example of the very, very exciting company that I think lots of people are excited about. It's not an insight of ours that this company is exciting, but getting them to add viruses and not just cancer to the things they're

trying to diagnose, because of our view on biosecurity. We've made investment in Impossible Foods. So a lot of the stuff we do in science ends up being just an extension of our other work. Like science, that's primarily interesting because of the causes that we have, but also as soon as we just see something outstanding and exciting, and a lot of times it takes the form of something that already looks good, but you could, with a lot of money, speed it up a little bit, and that would be worth it. And that might not be like the thing that other people are excited to do. So that's been a theme. Mm-hmm.

36:10 S?: Besides measuring animal consciousness, what are some moral philosophy questions for which better answers would change your decision-making?

36:16 HK: Yeah, what are moral philosophy questions, which better answers would change our decision-making? I mean, the biggest moral... There's like... I would say there's three categories of moral philosophy question that seemed really important. And believe me, I kind of wish it was not the case that there is any philosophy question that was really affecting our decisions. I think it's like a state you want to avoid being in if you can, in general, because I think that the methods of philosophy are not methods I particularly trust to get to greatly valid answers, but certainly the one you named. I mean, so you can help... Let's say hypothetically, that for $1,000, you can help one person a certain amount, or you could help 1,000 chickens a certain amount or 10,000 chickens a certain amount. How do you choose between those?

37:00 HK: And I think most people's initial answer is, "Why, I know how I choose, I just choose the person. What are you talking about?" And then the rejoinder might be, "Well, why is that? Do you think that chickens suffer the way that humans do? If you don't think that, why don't you think that? How confident are you?" And what we found is that, I think basically, anyone who really thinks about, "How do I know that humans are conscious and chickens aren't?" or "How do I know that human suffer more than chickens," ends up realizing that they're totally confused, and they have no idea why they think that, or even what any of these terms even mean. And so, that becomes kind of the domain of philosophy, and that is something we do have a write-up, it's public, it's called "The Report on Consciousness and Moral Patienthood".

37:42 HK: And I found it fascinating. It's kind of dry, it's very long, but I really recommend it. And I think it pulls together everything we can about animal behavior and about philosophy of mind, and tries to get to what is suffering, and how much of it is there in humans and chickens, and largely ends up throwing up [38:00] ____ but at least after reading it, I feel like I know what I don't know. And so, that's a moral question that... A philosophy question that affects us. Another one is like, how should you value future generations? So, if you were to reduce the odds of civilization ending or humans going extinct, should you think of that as being as good as saving eight billion lives, which is the number of people alive today, or should you think of that as the equivalent of saving 20 billion lives, because each of those people you saved could have had an average of two and half kids or something?

38:34 HK: Or should you think of it as the equivalent of saving a trillion, trillion lives, because that's how many people might actually just... Or what if it's a trillion, trillion, trillion, trillion, trillion, which some people actually think that actual number. And so, then you get into questions that are like... I think a lot of people have the intuition that increasing the probability of someone getting to exist is not morally the same as saving a life or preventing a death. And yet that position, which actually is my position, all things considered, but I have to admit, it's a hard position to hold up when you're engaging in thought experiments and trying to make it consistent, trying to make it a consistent, coherent position. And so, that has moved me a bit. And so, that is one of the questions we struggle with.

39:16 HK: And then the third question we struggle with is like, if you would spend your money one way depending on one philosophy view, and another way depending on another philosophy view, and you don't know which one is right, but whichever one you pick, it would swing all of the money, what do you do? And one aspect of that is there's a topic called moral uncertainty, which has a very small literature in philosophy. There's also a bunch of practical considerations about, do we want to be the kind of organization that just takes our best guess on Thursday and puts all of the money that we're putting behind that best guess, which says like, "Today, we think chickens are valuable, so we're not going to think about humans at all"? And I think there's a bunch of practical reasons not to be that kind of organization that we've written about in blog posts called "Worldview Diversification". So that question of like... It's kind of like a half-philosophical, half-practical question, but what do you do when a decision really seems to hang on some moral question that you feel extremely fuzzy and confused about, and how do you actually practically resolve that and make your bet? Yep.

40:19 S?: So, [40:21] ____ on that, you mentioned allocating funding into different philosophical buckets [40:27] ____,

and I'm curious how much you actually expect your [40:32] _____ final math funding to match those things given... You're currently giving about 100 million or above that [40:36] _____, and where do you expect a [40:39] _____ hit will be able to find ways of giving five billion to one of these buckets, whether in fact most of the [40:44] _____ will be determined by what thing you can find that can actually absorb that level of funding?

40:48 HK: Yeah, so the question is, we laid out our thoughts about a year ago in a blog post called "Update on Cause Prioritization at Open Philanthropy" on where we thought we were going to allocate the money in general, in a high level, like how much do we think we're going to allocate to animal causes that are really about helping animals, how much causes that are about helping humans, how much causes that are about reducing these risks in the long future, and the question is, how much we think we're going to stick to that, and how much we think that's going to get changed by practical considerations about where we can actually spend money? I think my answer's basically no idea. And I think we are working on ways... This is the cause prioritization challenge that I think we need a whole team for. We need to find ways to make the philosophy and the practical questions interact in a satisfying way, so that we can make informed decisions and decide in the end where the money's going to go. And I think at the end, we may and end up doing a significant amount of throwing up our hands and making arbitrary decisions, we want to make them in as informed a way as possible.

41:50 HK: But one tool that we're using right now is we just say, we'll make a table, and the rows of the table will be different causes that we could put more money into, and the columns, each column will be like, how much good per dollar would we accomplish if we expanded this cause, and it'll be according to a different philosophical view. And so, you can look at this one table and you can say, "Okay, if I give more to GiveWell top charities, then I do this much good under this philosophical view, and that much good under that philosophical view, whereas if I give to farm animal welfare, then I do zero good under the human-centric view and a lot more good than GiveWell under the animal-inclusive view." And so, this is like we started to build tables like this that kind of roll all the considerations together at once, and it's not about solving the problems, because we don't think that we're going to solve these philosophy problems, but it's about having a really informed choice to make and creating something that ultimately is just going to inform Cari and Dustin and help them decide where they want their money to go, and then any other donor we advise, where they want their money to go.

42:52 HK: And so, that table does tend to kind of jam everything together, and it also includes the question you're asking. The rules will also include how much money could we put into this, that we know of? What's the most money we think we could put into that cause and get this kind of return? And if we put more money in, the return would fall, and so, how would that change? And so, that table is going to evolve every year, and every time that table evolves, you're just going to look at the table and it's going to say something like, "Well, if I try to put it all into long-term stuff or global catastrophic risk stuff, then the first three billion would be really good, and maybe the next like billion would be not as good, and whereas if I put it all into animal stuff, this is what I would get. And if I split it up, that's what I get." And you just have to look at these trade-offs and say, "Where do I want to spend the money given what I can get for it?" And so, a lot of this work involves a lot of philosophy because we have to identify the key questions, but at the end, we're just trying to make sure people are making informed decisions because it is philanthropy and there's not going to be like an objective right answer. Yeah, you... Sorry, in the back, yeah.

43:54 S?: Is there any thought to having a method for smaller donors, not multibillion dollar donors, to do hits-based investing?

44:08 HK: Yeah. Is there a method for smaller donors to do hits-based investing? Certainly, you could just give the things that you think are really risky and could be really huge, and... But I think it's not... Advising smaller donors is not Open Phil's priority right now. And so, we do put out usually, I haven't decided if we're doing it this year, we probably will, put out an annual post that just says, "Here's where a bunch of Open Phil staff think you can give if you're an individual donor." But I have to say, it's not what we think about day-to-day because when we put together that post, everyone has to think about, "Okay, what's an organization I know that you could just drop a check on them, and the check could be any size, and they can then use that in a productive way?" And that's just not usually the way we think. Usually the way we think is like, "Alright, here's someone who needs $50,000, here's someone who needs $100,000," and we're always trying to make the grants the right size, and we're always trying to fund what we're excited about.

45:05 HK: And so, usually we have something, but it would be a different job if we were to optimize that way, and I do think... I will also say, I think hits-based giving makes the most sense if you either have a whole lifetime or all your time to build your own staff and your own networks, and have a lot more to go on than... Have your own kind of trusted people who can therefore make you willing to bet on things that no one else will, or you don't have the time to build

your own team if you just have people you trust a lot. That's when I think hits-based giving makes sense. I think I don't see a great fit between the hits-based giving philosophy and trying to be someone who thinks about giving for a few hours a year, and just tries to give to something that you can find from a public source. So I think that's just, would be a longer conversation why I think that, but that is why we... It's gone hand-in-hand a bit, the hits-based giving philosophy and the focus on the kind of mega donors. I don't think those two always have to go together, but we have tended to focus on the mega donors, and we tended to focus on hits-based giving, and it just, in our attempt to stay focused, that's been where we've mostly stuck to, when we do that annual post. Alright, now you.

46:19 S?: So your answer to the science-based question is the direct non-meta answer about science projects that can directly do good. Your kind of analogy to the philosophy question's question, do you spend much or any time thinking about what scientific, what questions or areas could produce information that would help you figure out other things in order to direct the giving better?

46:44 HK: Sure. So do we think about what grants we can make that would help us then gain information and do giving better? We do think about it. I wouldn't say that it's like a cause, I would say just that we're always trying to figure stuff out, and we always try to have that muscle that says, "Do we think we could fund someone to figure this out instead of us doing it ourselves?" So, we've definitely done a fair amount of things like that. An example is, we do work on potential risk for advanced AI and other causes where there's a big debate kind of in the background about is it possible to predict the future at all? Is it possible to say anything meaningful about the future? And as part of that, we did make a major contract to an outside firm that's collecting all these predictions from past sources that have made long-term predictions and kind of scoring them and seeing how they went. So hopefully we'll have that out sometime in the coming year.

47:36 HK: So it's generally like, yeah, when we think of it, when we think we can make a grant to someone in a way that will inform us and leave us in a better position... I know a lot of animal work has just been funding research so that we can figure out what would be good standards to improve animal welfare. It's not always as simple as get them out of a cage. Like for broiler chickens, or for fish, there's these questions like, "What would be the easy interventions that would make the animal suffer less?" So we funded a lot of that, and that's been a big interest for the animal folks, but it's usually on as needed basis.

48:08 S?: And the sub-category version of this question is the cross or now that there's a split, GiveWell will never fund the people who are running the RCTs that they depend on.

48:18 HK: Well, they do.

48:19 S?: They do?

48:20 HK: Yeah, GiveWell totally does that. Did you talk about GiveWell Incubation Grants?

48:24 S?: No.

48:24 S?: No.

48:24 HK: Okay, well, yeah.

48:25 S?: Great. So, because they do it themselves but you don't see a role for Open Phil to try to give to the science that GiveWell uses, instead it's more [48:35] _____ depending which organization.

48:35 HK: Yeah, Open Phil general... So GiveWell has this program called the Incubation Grants, which is more like they're higher-risk grants and the goal is to create more top charities. And actually, the way the process works right now is GiveWell does a write-up, and since we're the primary advisor of Cari and Dustin, and they're the primary funder of the Incubation Grants, we are part of that process for the Incubation Grants. But we basically just do whatever GiveWell says, and that's our main process on Incubation Grants. And that's how we... We're kind of formally part of that, I see all those write-ups, but that's really led by GiveWell. And we also make grants that inform our work in other areas. So, the short answer to your question is like we are totally good with supporting grants that give more information about how to do our work better. In GiveWell's case, that's GiveWell Incubation Grants and they decide how to make those, in our case, that tends to be about the causes that we work on. Yep.

49:34 S?: So, GiveWell and Open Phil can kind of yield better than market returns because the markets are not exploitable. This is kind of a big question, but is there any work, or do you know if people are doing things to make markets exploitable so that things will organically happen?

49:51 HK: So I think the question is we're... We believe, we hope that we're doing things that are especially good, and in some sense, beating the market. In some sense, if you think about a market as a world where every time something is good, someone does it, we think we're kind of breaking that rule by doing things that are good that people aren't doing. And have we thought about ways to make the market efficient, which I guess I'd interpret that as, have we thought about ways to change the world, such that good things will just quickly and automatically get funded by someone and then you don't have to have this special process? And I guess the only way I can think of to do that would be to convince more people and agencies that have money to spend the money the way that we think is the best way to spend it, or to use the same kind of principles as we do, accomplishing the most good for it. And I guess my answer to that would be like, that is something that we are interested in doing at some point, but just hasn't been our top priority because I largely want a better... I want Open Phil to be more mature and more developed and more solid in its own thinking before my top priority becomes convincing other people to do what it's doing. Does that make sense? Cool. I'll take one more, maybe two more. Yeah.

51:09 S?: You mentioned briefly immigration reform as a cause area, can you talk more about that, and also if you have other cause areas that you don't have a program officer for that you'd want one for?

51:18 HK: Yeah. So, there's a bunch of cause areas that we don't have a full-time program officer, and we'll either fund the occasional outstanding grant or we're just looking for one. So the immigration reform is one where this is... If you look at a lot of the development economists who are kind of in the world around GiveWell, a lot of them will point out, "Actually, the best anti-poverty intervention in existence that we know of is immigration." When someone steps across the line from a poor country into a rich country, it's like their income [51:48] _____ might go up by 10X, and there's no other intervention that does that. And usually in immigration debates, it's like one side is saying, "This is bad for America," then one side is saying, "This is good for America," then occasionally, there's a crazy third side saying, "Well, what about the brain drain, what is this doing to the host country?" And the thing that comes at the least is like, "What about the immigrant themselves? Is this good for the immigrant themselves?" And in fact, it's so good, we believe often for the immigrants themselves that that should be the dominant factor.

52:16 HK: And so, we just think immigration policy is a great example of a policy area that has enormous stakes, like really high in the importance scale. That is something that we got interested several years ago. I think the tractability looked really bad then, and now it looks atrocious. And so, it just... [laughter] It looks worse than it did, and so it's not an area we're incredibly active, although we are working on a grant right now to an organization that we are actually excited about, that might help advise bilateral immigration agreements, so countries making deals of each other for mutual benefit that I think may side-step a lot of the politics of what's going on in the main immigration debates. So that's a cause we're interested in. We're interested in macroeconomic stabilization policy, which is a very wonky cause. Maybe many of you haven't heard of it, but basically the Federal Reserve, the Central Bank, also the Central Bank of Europe, in Japan, UK, they make these decisions about what sort of policy to run with respect to the macroeconomy, and it's generally perceived to be this very dry technocratic apolitical thing that no one should worry about because the economists all have it taken care of. Our view is this is probably the single biggest lever for the fortunes of working class people, and it deserves to be thought of more through a humanitarian lens than it is.

53:40 HK: So we think that actually the current culture of these central banks is a little bit overly obsessed with inflation as the thing they're trying to avoid and not... I think if they weighed just like what is the benefit to humanity of having one mistake versus another mistake, we think they would weigh them very differently from how they do now. So that's a pretty hairy cause, and it's one that has been very hard to find a great program officer for, but it's definitely one that we're interested in. So those are just a couple of examples. I'm not going to... Really, a lot of Open Phil is just bottleneck by not having all the staff and understanding and time that we would to do everything that we could. So there's many, many causes we would be interested in if we had someone who we found who was amazing and who was ready to do the work that we need done in them. I'll take one more. Yeah.

54:35 S?: I feel like some of the parts of Open Phil are fairly traditional grant-making in [54:39] _____ and doing that sort of thing, and somehow seem much more novel, like moral patienthood [54:45] _____ and so on. Is that like... Do you have explicit frames for like, how does parts of the org interact with each other? Do you think of them as distinct, or are you like, I'm just doing a single org and it's all, it makes sense as a whole?

55:00 HK: Sure. I don't think I have much answer to that. The question is like Open Phil does some things that seem like things other foundations do, like making grants than it does things that don't seem like things other foundations do, like this cause prioritization project where we're thinking about what does it mean to suffer, and blah, blah, blah. I don't have a great answer. I think it's an empirical fact that the more normal in activity, the easier it is to hire someone for that activity, and the more weird in activity, the more we tend to go to our generalists, and the more we tend to just say, "Whoever it is who's flexible and knows Open Phil really well will do this job." And that affects how we allocate priorities. A lot of times it... Yeah, a lot of times, it's easier for us to do things that are less unusual because it's easier to find people who are ready to come in and do them. So that's probably the main way in which it's a noticeable difference, but I think mostly we just try to prioritize things according to how much they'll do for a mission as a whole.

55:50 HK: Cool. So, I'm going to wrap up. So thanks for all the questions. It's been a lot of valued questions, and thanks for all your support.