

**San Francisco Research Event, June 14, 2016 – 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 Speaker 1: He kind of told you all the rules, that if you want something taken out, just let me know, either as you say it or afterward. But thanks everyone for coming, I'm Holden Karnofsky, I'm one of the co-founders of GiveWell, Elie is the other one. I'm gonna talk about the Open Philanthropy Project which is still part of GiveWell.

00:16 S1: We are hoping that some time in the near future, maybe a year from now, it'll be a separate organization, because we think that it's getting distinct enough and the missions of the teams are getting distinct enough that that's going to make sense. But I'm gonna go over the basics of what we do and give one example of the kinds of causes we're working on, and then I'll take a break and questions, and then talk about some of our other work and then take more questions. And I'm gonna shoot for around 9:00 or 9:05 to end this.

00:45 S1: Basically, Open Philanthropy Project, for those of you who aren't familiar, it has something in common with GiveWell, in that we're looking for the best ways to accomplish good by spending money. And it started in 2011 when Elie and I met Cari Tuna and Dustin Moskovitz. They were looking for the best ways possible to do philanthropy, and it reminded us of a problem that we had had when we were starting GiveWell. Except instead of, Elie and I when we started GiveWell had been trying to give away a few thousand dollars and we had a few hours a year to think about it. And Cari and Dustin were trying to give away a large fortune and had several decades to think about it. So those are kind of the similarities and the differences right there.

01:27 S1: So, Open Philanthropy Project is more targeted at good ventures... Sorry is more targeted at major donors such as Good Ventures, which is Cari and Dustin's foundation. And the basic idea is that where GiveWell looks for proven, cost-effective scalable charities, Open Philanthropy Project looks for important, neglected, tractable causes. And we try to find the best causes to work on, and then we try to hire around those causes, build staff capacity, build knowledge, build expertise, and become really good at working on the causes we've chosen, and then make grants in those causes. And that takes advantage of the larger donor's ability to build expertise, have more engagement, have more context, and also do things that aren't just writing checks to existing charities. So that's why their kind of unit is the cause instead of the charity.

02:18 S1: One of the very distinctive factors about what I call Open Phil as opposed to GiveWell, is the emphasis on what we call hits-based giving. So we had a blog post laying out this idea a couple of months ago, but the basic idea is that when you're doing high-risk giving, you may do a whole bunch of... You might support a whole bunch of things. And maybe for every 10 things you support, nine of them completely fail and have no impact and one of them has so much impact that it makes up for everything else combined.

02:49 S1: This is kind of conventional wisdom about how venture capital investing often works, and I think there's at least some flavors of philanthropy that I expect to work this way as well. So when you look back in history at some of the biggest philanthropic success stories, and we do look back at history a fair amount. We have a project where we try to study the big successes philanthropy has had, you do see some that look so big, even though they were very risky at the time, that you could imagine a philanthropy having a bunch of failed projects and having this one success that makes up for everything else.

03:22 S1: So the research that the Rockefeller Foundation did that has been credited for just getting maybe more than a billion people out of starvation is one example of that, where it was high risk, agricultural research, nobody really knew where it was going, but you could have a lot of failures like that. You could have a lot of failures if you have one success of that magnitude.

03:42 S1: And so that's what we're about. We do things other than hit-based giving, but I bring up hit-based giving to just give some context on some of the things we do and why they can appear a lot more high risk than what GiveWell does. And often times are a lot more reliant on kind of the expertise of the networks we've formed and the people we've hired and a lot of deep contexts that we can't fully put into writing and justify with evidence the way we do with GiveWell, though we do try to write up enough of our thinking to help people follow along with the basics and see what our basic approach is.

04:20 S1: So in terms of what we've done and our progress, for our first few years we were really focused on picking the causes we wanted to focus on. So finding good areas for a major philanthropist, finding things that were important,

neglected and tractable. And in 2015, we had picked a bunch of causes to focus on and we spent a huge amount of the year just hiring, and building capacity and trying to get to the point where we had enough staff to really make a lot of grants.

04:49 S1: And in 2016, we have sort of got twice the size that we had at the beginning of 2015, and we do have a lot of specialized staff and we do have a lot of focus areas that we're ready to make grants in. So for 2016, a much bigger focus has been actually getting money out the door and making grants, and that's some of what I'm gonna talk about.

05:08 S1: I'm gonna start by just going over the basic idea of our US policy work and some of the work that we've done there, and that should illustrate the basic principles we're using and the basic things we're looking for. And then I'll take questions for a while before talking about some of the other things we're up to.

05:23 S1: So, a couple years ago we started looking at the different ways we could get involved in policy-oriented philanthropy. And so, that is giving with the aim of envisioning a world with better public policy. And the idea is that, if you were to support someone who is working to educate people about how policy could be better, or do some of the grunt work needed to improve policy, you might spend a relatively small amount. But if you end up having an impact on policy, that can leverage your giving greatly. And so, this is one category of thing that we think could be very high impact. And what we did is, we took a whole bunch of issues we could imagine working on, and for purely pragmatic reasons we kind of started with the US universe 'cause that's where we were able to understand the issues and form our networks. Although as we've gotten into specific areas, we've become more open to international work, and I'll get to that.

06:18 S1: But we made a list of all the issues we could look at, and we tried to figure out which ones... We took a lot of issues and estimated their importance in terms of, if we could get a big policy win, how much value would that create for a society? We looked at their neglected-ness. In other words, we prefer to work on causes where there isn't already a lot of money and a lot of philanthropy and a lot of groups trying to push on the issues. So, some issues we think are quite important and quite worthy we think get plenty of attention already, and other things we think are equally worthy and get less attention.

06:50 S1: And then tractable, we do pay attention to what we think the prospects are for actually winning some victories. Although I should distinguish that because as a philanthropic organization, we operate on long time horizons. And so, when it comes to policy, for us, tractability does not mean there's a big story about this in the news right now. If we move in the next three weeks we might be able to make a difference. Because we operate on a scale of supporting organizations as they grow, and that timeframe is just not a timeframe we can operate on.

07:22 S1: And in general if something's all over the news, the people who've been working on it for a decade are the people who are now gonna make a difference. And the people who try to jump in now that it's in the news, are probably at high risk of wasting quite a bit of money and effort. Relative to the other people anyway. So what tractable means to us, is something more like, there are deep structural factors that we expect to persist for a while that make it possible to get wins on an issue.

07:50 S1: And so, an example of an issue that we have been drawn to for tractability reasons is, criminal justice reform. So, we believe that... Well we know that America incarcerates more people than any other country on Earth per citizen. We believe that that incarceration does not have a commensurate benefit in terms of public safety. There's little evidence that it does. David Roodman is currently conducting a very massive review of the literature on this, and concluding that there's not a lot of reason to expect that with the current levels of incarceration that we're really getting much for that extra incarceration in terms of public safety. So we do incarcerate a lot of people. There's questions about whether that's actually paying off. And so, we believe that there's a lot of humanitarian gain from sort of rationalizing criminal justice policy. But if that were all there were to the issue, there'd be a lot of other issues that are competitive with it in our view. One of the things that has really made it stand out is, the idea that there is momentum and opportunity to make change in this area in a way that a lot of other issues don't have. And there's multiple reasons for this.

09:02 S1: One is that, it's a state and local issue. And so, you can pick your battles a lot more than with issues where the only venue you could ever possibly, sort of make a difference in, is DC especially with the way that things work right now with Congress and the situation in DC. So, another reason is just, crime is a lot down. Since the tough on crime days a lot of state budgets are tight. There's some rare opportunities for liberals and conservatives to come together and say, "We have opportunities to help the most disadvantaged citizens by cutting the size of government." You don't see

that all the time. And so, that is an example of a cause that we were drawn to because, we estimated it had high importance, though not as high as some of the other causes we looked at. And that it was outstanding in terms of the tractability. And our first major hire was in this area. So we hired Chloe Cockburn, who comes to us from the ACLU. She's been with us for about six to nine months. She joined in September. So I'm not very good in math, but it'd be about nine months. And she basically leads our work on criminal justice reform. She works out of New York, she's not here tonight.

10:11 S1: But just as an example of one group that we've supported here, The Alliance for Safety and Justice is a group that has grown out of Californians for Safety and Justice. Californians for Safety and Justice is the group that led the campaign on Prop 47 in California that reclassified a lot of felonies as misdemeanors, and therefore in our opinion reduced the amount of time people are gonna serve in jail and prison in positive ways. It was a really well run campaign. And now the people who did it are trying to bring their skills to the national stage via The Alliance for Safety and Justice. And we believe based on our relationship with Chloe and her relationship with people in the field, that these are some of the best people in the field. That there are some of the best position to do a lot with a lot of resources. And we've given them substantial unrestricted support. So there was a \$2 million grant that was announced and that's an example of the kind of thing we're doing where it's not... We didn't take this group and find 10 randomized controlled trials for the work that their doing to change policy. It was much more we picked the cause, we picked the people based on networks, and we looked for someone who could use money and who needed money, and who had that kind of great reputation and great track record in a softer way. So that's criminal justice reform.

11:27 S1: Other example and then I'll take questions. But Farm Animal Welfare is a cause that we put high on our list because the importance, depending on exactly what your philosophical views are on how to sort of weigh the suffering of animals, but the importance could be very high. There's a lot of animals that are treated very, very badly on factory farms right now. And we also felt that this cause was relatively neglected. We felt neutral on tractability, we weren't sure what to think. At the time that we made the hire. But now Lewis Bollard is leading our work on that, and I don't know if he's here tonight, but he's in town right now... There he is.

12:08 S1: And so, when we started our hiring process Lewis came to us and actually told us some things we didn't know at the time, which is that there actually has been big signs attractability and there's been a lot of momentum around cage-free reforms. And so, around lobbying corporations to make pledges that they're only going to sell eggs that are cage-free, and they started with fast food and moved on to grocers. We believe that cage-free eggs, they're not perfect, we still don't like the way chickens are treated under cage-free systems, but we believe they're a lot better, and we also believe that it's a really good thing for this movement to get some wins. We think that's going to be good for the longterm growth of this brand of animal welfare and for the longterm prospects.

12:50 S1: And so, we made a series of grants, the Humane League and Humane Society, and one other... Mercy for Animals, yeah. That was our first set of grants was to support the work of three groups doing this and get them to do it in as accelerated a manner and as ambitious a manner as possible, and shortly after we made our recommendations. We believe that we helped them plan these efforts in a more ambitious and speedy way, and they have recently been winning quite a few victories and now it looks almost inevitable that the US is gonna be almost a cage-free country, in terms of groceries and in terms of fast food, and that's going to affect a huge number of chickens and more generally be just a big win of the kind that we think is important. So, those are examples of some of the stuff we've done so far. I'm gonna pause and take questions before talking about the rest of the work that we're doing, but that's the basic idea. Questions? Yep?

13:50 Lewis: Can you just real quick explain what the divisions are? Like if I give money to GiveWell, is it going to the Open Philanthropy Project?

13:58 S1: Sure.

14:00 Lewis: And what's the basic relationship?

14:02 S1: So relationship between GiveWell and Open Philanthropy Project and what happens if you give money to GiveWell. So, right now Open Philanthropy Project is just part of GiveWell, it's in the same organization, and if you give to GiveWell unrestricted, that means that you are supporting both, it's going into one pot of money. That is something we're looking to change. We are looking for the two organizations to become separate, but that could be a pretty involved process and it could take a long time. And if you want to support GiveWell without supporting the Open Philanthropy Project you can certainly say so when you donate. Also if you give to GiveWell and you restrict your

donation for grants it will go to top charities, at that point it's not really part of the Open Philanthropy Project, but that is the kind of situation. Yep?

14:49 Speaker 3: You said that you think it's pretty likely that the cage-free movement will end up succeeding at this point. To what extent do you think that was helped by organizations like Humane Society, Mercy for Animals, and to what extent were those organizations in turn helped by GiveWell?

15:04 S1: Sure. Sure. So, what's the impact of the groups we supported on cage-free wins and what's the impact of Open Philanthropy on those groups? So the first one we feel like the answer is huge and essential, and I'm gonna... Maybe Lewis you should jump in if I'm wrong about this or maybe you should answer. Lewis certainly knows a lot more about this than I do, and all I'm doing is repeating stuff he's told me. But we definitely think that those groups, we just don't really see any reason to have expected any of these pledges to happen without the work they were doing. We believe that there was no reason to think grocers and fast-food companies were just gonna decide to go cage-free. They were under pressure to do so, they were being lobbied to do so. So we think... And we don't know of other groups doing this and we believe we would know if they existed. So we think that's just an essential connection and basically you can put 100% of the impact there.

15:57 S1: In terms of our impact on the groups, I think that one is dicier. They already had some wins when we started talking about supporting them, and all I can... All we tried to do was speed them up, make them more ambitious and faster. I think there's definitely an argument that we sped them up a noticeable amount. I think there's even an argument that if they hadn't moved quickly, that there could have been a lost opportunity because some of these companies were investing in these enriched-cage systems, and once the capital investment is there, that's gonna be a harder pledge to get them to make. So there's some possibility of that, but I would say in this case, I would say that it looks like a pretty strong case that we made things faster and more ambitious. I don't think that we were the only factor there at all or a primary factor, I think most of the credit goes to those groups. Yep?

16:54 Speaker 4: So, I also recently heard about another victory the Humane League has won in terms of eliminating chick culling. Is that something that Open Phil had a voice in, or was that separate for the work that you were an accomplice [17:05] ____ for?

17:06 S1: So the question about the Humane League recent victory on chicken culling, how involved were we in that? Look, one of our philosophies as a funder is we are not generally doing contracts, we're generally doing grants. And so, what we did do with these groups is, we made the grants and we made it clear that we were trying to support corporate campaigning. I don't think that we said, "This is only for cage-free." I think that we said, "This is for corporate campaigning, this is what we think is the most promising branch of Farm Animal Welfare work right now."

17:40 S1: And so, I think this is a tough question to answer in the sense that probably they worked on this chicken culling thing because they had won victories on cage-free and were looking for the next thing, and so, I think in that sense it's connected, and I think that our funds are bound up in that work. It isn't something that we asked them to do, it's also not something that we asked them not to do. It was more like an example of the kind of thing we were really happy that they did and were trying to support. Does that sound right, Lewis?

18:08 Lewis: Yeah, I would say even more than that [18:11] _____. It was the cage-free campaigns that brought The United A producers to the table. With the Humane League they were very explicit that they were only meeting with the Humane League because they had so much success with the grocers. And that's what allowed the Humane League to credibly say they would campaign on this issue [18:30] ____ without any campaigns, succeed on this issue, was that they'd shown with their previous campaigns on cage-free [18:36] _____. And then it comes back to the question of what role we had in those campaigns. And with mainly this case, we tripled the size of their campaign team, but they still would've had campaigning without us, so you have to guess what they [18:50] _____.

18:52 S1: One of the things you're observing here is something different about Open Phil and GiveWell, is that knowledge is more distributed, and so...

18:58 S5: Well, Elie had Josh and Natalie talk about their [19:01] _____.

19:01 S1: Yeah, true. Okay, fair point. The thing I'm trying to say is that we do... My job at Open Phil is to work really hard on picking the right causes and hiring the right people for those causes. And then the gulf between what they know and I know is quite extreme. And I think it's bigger than for GiveWell, though with GiveWell stuff that's in progress is

more that way. But I think with GiveWell there's much more of a mission to take the conclusions we're reaching and really write them up, so we can cross the full bridge from not knowing anything to understanding the full case.

19:37 S1: And with Open Phil there's much more of a sense that we're hiring people who have deep context, who have extensive networks, and I can't keep up with everything they know and everything they're doing. And so, the goal is to be really good at picking the causes, picking the people, and then we have a lot of interaction in terms of asking questions about what they're doing, but it is fairly high-level and a lot of the philosophy here is to trust the program officers to lead the way and to be in the lead on setting those priorities. They do get questioned a lot, but it's a different kind of basic model. Other questions? Yeah?

20:14 S3: Do you feel the need to see the causal connection between the giving and the impact, or do you not feel that need?

20:21 S1: Do I feel the need to see the connection, the causal connection between the giving and the impact? I would say yes and no. I would say that when a grant proposal comes into my inbox, I'm asking about causal connections, forward looking and probabilistically. A grant that makes me very excited is when I can say, "I think this has a decent chance, maybe 10%, maybe 20%, so not necessarily likely, but a decent chance of making a really big difference to something really important." And if you multiply that through, this is a great value for money in expectation.

20:56 S1: And when I'm thinking about whether something has a decent chance at succeeding, I like to play out causal stories in my head and to imagine that... I can imagine what's going to happen, or at least examples of what's going to happen, that's going to result in good things. Though, there are multiple models of grants. And so, I think there are some grants where we really just believe in the people or the institutions we're funding. We don't know exactly what they're gonna do. We know they've done good stuff in the past, and we're just trusting them, and figuring that we'll come back in two years and we'll see something good and we don't know what.

21:27 S1: And that's the kind of grant that we make. And I could name several grants like that. I think ASJ in large part is like that, the Alliance for Safety and Justice. I think the corporate campaigns subject to being corporate campaigns are like that, as well. So I think that's definitely a model. And then there's other models where, let's say that we know what's going to happen and we think that that thing that's gonna happen is good. So we do both.

21:50 S1: I like to be able to have examples of causal impact, either in the past or just examples of good things that could happen. And then in terms of watching for results, I think there is an attitude... The attitude we try and take is that a lot of these grants may be operating on very long time horizons and may be very hard to isolate the impact. And so, we have to be patient and willing to do things that look good forward-looking, and just have a very long time horizon and a big portfolio for looking back and seeing how we've done.

22:20 S1: I think the corporate campaigns have been successful more and faster than I expected, but I think that's the exception. And there have been some other wins that I would say have already been exciting, but for the most part we expect to learn how we're doing more by looking across the whole portfolio after 10 years, than by taking each grant and marking off milestones. Does that make sense? Cool. Other questions?

22:48 S1: Why don't I talk about some other causes we're doing. So, I talked a little bit about the policy influence category. Another category for us is global catastrophic risks. We think a very good fit for philanthropy is working on worst case scenarios because if you imagine that... You imagine something happening, like a pandemic, the very, very worst pandemics imaginable are way more damaging than the only really bad pandemics, but they're not necessarily way more unlikely, even though they're way more damaging. And so, I think there's a general feeling about global catastrophic risks that, the things that could be so damaging as to destabilize civilization in general are not the kind of things that any other institution really has the incentive to work on, other than a philanthropist. And that we might be able to have an outsized impact by improving society's preparation for things that could have global and long-run consequences by being very disruptive.

23:52 S1: A couple years ago, we made a list of all the things that we could imagine going wrong and being very disruptive to global civilization. We had asteroids on there, and climate change, and antibiotic resistance and a whole bunch of other stuff, and we did the same thing. We rated them by importance, neglected-ness and tractability, and ended up with two main focus areas, that I would say are the two things that I think are probably the most likely to have a really globally de-stabilizing, threatening human extinction, or just a real change to long term trajectory of civilization.

24:28 S1: The two causes that I think are the most likely to be in that category. One of them is pandemic preparedness and biosecurity. So, I actually do think that whether a natural pandemic, which has been behind some of the biggest tragedies globally to-date, or a synthetic pandemic as synthetic biology advances, if I were trying to come up with things that could really knock civilization of course in the next 100 years, that would be near the top of my list. And then the other cause that we've made a focus area is potential risk for advanced artificial intelligence.

24:58 S1: This one obviously is a lot less intuitive to immediately see the case for, and I myself have a history of changing my mind a lot on this issue. It's a long thing to get into. I'm not gonna recite the whole set of blog posts that we put up to explain our issue, to explain our interest in this cause, but I would broadly say that if and when, and I think no one really knows if and when, but I think, if and when we get artificial intelligence that is powerful enough, that can do enough things at a super human level, that is one of the top candidates for just massive changes across the globe because, most of the changes you've seen across the globe to-date can be attributed to human intelligence. And something that was noticeably more powerful than human intelligence in a large number of relative domains, for good or for ill, could be one of the most transformative things that has ever happened.

25:52 S1: And we believe that there's a series of reasons to believe, and a series of potential risks that if things go a certain way, you can imagine either very catastrophic accidents where you have a powerful artificial intelligence trying in some sense, optimizing for an outcome that's not human friendly or misuse. I think if you imagine that powerful enough artificial intelligence would be a very powerful weapon, that weapon in the wrong hands could be globally catastrophic. And so, we believe this is an area that is definitely speculative. I know it sounds a little science fiction-y, but we also think it's real.

26:30 S1: We think there is a real chance that something very transformative is going to happen via AI in the next 20 years or so. Lots of reasons why we think that, written up in great detail. Well, not great detail actually, but written up in some detail on the blog, and we believe that it's a pretty neglected space. We believe that it's partly because it's a little bit out there. These set of people trying to prepare for these unlikely but very high stakes risk is not as big as it could be, and we believe that as a philanthropist, we can help build certain fields. So, those are the areas we're working on.

27:05 S1: I should mention climate change, which would also be the third thing on my short list of things that could really be huge. I think climate change is one of the biggest risks facing humanity. I also think it's a lot less neglected, especially by philanthropy than the other two. And I think the very worst cases are probably not as likely to be as bad as with the other two. And that would be a long conversation too, but we have those two causes and on biosecurity and pandemic preparedness, we have Jaime Yassif and Howie Lempel working on that. And they're both focused on that cause, and Jamie joined us recently.

27:43 S1: We've made only a few grants so far, but I expect that in a few months we'll have a lot more to say about all the work we're doing trying to improve pandemic preparedness across the globe. And then Elie Hassenfeld is actually managing that work. So that is probably... Out of all the causes I'm saying tonight, that's gonna be the one I know the least about [chuckle] because Elie is doing the supervisor role with that particular work, and then on potential risk for advanced AI, that is something where Daniel Dewey is working for us full-time. He was a contractor last year and he converted to full-time this year. And then also several generalists are spending time on this cause, including myself, partly because we think this is a very important cause and a big opportunity at this stage. Especially as interest in the issues is growing and we think those are opportunities to get things done really well or really badly.

28:35 S1: And also just because it's a particularly young cause where the networks are not very well developed and the field is not very well developed. So, it feels right now very time intensive and it feels like we need to find staff time where we can find it, if we're gonna position ourselves to do the kind of work in this cause going forward, that I think we are already able to do on criminal justice reform and pandemic preparedness and farm animal welfare.

29:00 S1: Those are other causes we're working on and I'll just quickly hit other stuff that we're doing because we're not done picking causes. We've recently had several people or will have several people come on, to help us explore the world of scientific research, which is one of the hardest categories to explore in terms of things we might recommend funding to, that could be really good value for money. Daniel Martin-Alarcon is here. He recently joined us as a full time science advisor. We have two other people coming on full time in the next couple of months. And so, that's a cause that's taken us a long time just to get the staff through the cause selection, but eventually, hopefully we'll have some focus areas. And then, there's other work we do. We have the history of philanthropy project, but I'll leave it at that for now and spend the rest of the time taking questions.

29:54 S5: Jamie's here too. Hello?

29:58 S1: Yep. Sorry, I didn't see you.

30:03 S6: What's going on with history and philanthropy?

30:04 S1: What's going on with history and philanthropy? So, history and philanthropy is basically... We have... We've gone... We tried to go through all the books we could find and all the literature we could find on what philanthropists have accomplished in the past, and what we can learn from what they've done. And we have a big internal list and actually an external list too. It's called "Philanthropy's success stories," of candidate cases where it is claimed that philanthropy made a big difference. We have started the process of really deep diving those cases by actually contracting with historians, who then go try and check things out. So, for example, supposedly philanthropy played a role in the passage of ObamaCare, the Affordable Care Act. Is that really true? What were the key organizations? What role did they play? Who funded them? How did they decide to fund them? Ben Soskis, a historian that we work with, wrote up a really good case on that. I really like and it's public and it's on our history of philanthropy page. We blogged it.

31:02 S1: Other cases that we've been interested in, and that are coming out either recently or very soon, there's two organizations that we not only have found to make pretty good cases that they've had a lot of impact, but also that we believe philanthropists had a major role in founding not just supporting. And so, there's the Center On Budget and Policy Priorities, which is a group that works on pro-poor, poverty reducing work in the US and they try to be part of budget debates on Capitol Hill. We believe they've had a lot of impact and really had some high leverage wins. And then there's a group called Center for Global Development that we've supported significantly. It's an example of a group that actually we've supported them without really having a set of deliverables. And they're a group that works on how to make rich country policy more beneficial for the global poor. And they also... Both of these groups we believe are... At one time there was a philanthropist who said, "I believe there should be an organization like this" and they went and found the person they thought would be best to lead it, and then they funded that person and trusted them. They certainly didn't run things, they certainly didn't design things but they did initiate things, and we think that's really interesting 'cause it's counterintuitive to me that that would ever work.

32:19 S1: I think a week or two ago we published Suzanne Kahn's study of the Center on Budget and Policy Priorities and how that got started and then tomorrow we're gonna publish... Probably tomorrow, we're gonna publish Ben Soskis' study of how the Center for Global Development got started. So this is basically the work, we have this list of things that look interesting, we have historians check them out, we publish the case studies. It's not something we spend a huge amount of time on, it's something we do with contractors and we get the case studies done as they get done. And then there was also recently a new book that came out that lists an even larger number of success stories that we're gonna try and go through by the end of the year. It's called "The Almanac of American Philanthropy." So, it's not a huge use of our time but it is us trying to learn what we can from history and especially really closely scrutinize the successes and say, "Were these real successes and what can we learn from them?" Yep?

33:15 S3: How do you think about the neglectedness of strong AI risk in terms of the new interest in philanthropy around it in Silicon Valley, especially Musk and [33:27] ____ Wasi?

33:28 S1: Yep. Yeah. So how neglected is the field of potential risk for advanced AI? This is something we've looked into a lot and had a lot of conversations about. One thing that is true is that over the last couple of years, there's been a big jump in interest, I would say in AI generally in the general prospect that very powerful artificial intelligence that is superhuman performance in a wide variety of domains could be coming sooner than people had previously thought. And I think that is due to technical advances that just took a lot of people by surprise, and especially deep learning is a class of methods that has shown to have really exciting and broadly applicable effects. So, there's been a lot of excitement about AI and with that, there's been a surge in interest about risks for advanced AI and certainly a lot of press and stories with Terminator pictures.

34:23 S1: We believe that despite all these, it hasn't really translated to a lot of philanthropy and it hasn't really translated to a huge amount in the way of what we really want to see, which is people working full time to do things that could reduce potential risks. To us, stories in the newspaper don't reduce potential risks, hype doesn't reduce potential risks but people working full time do, and we think the amount of those has gone up, but it hasn't gone up a ton. And the amount of philanthropy has gone up but it hasn't gone up a ton. Specifically I mean OpenAI, which I think that the

groups you're naming, that's their main project.

34:58 S1: OpenAI is a non-profit research lab focused on... Basically focused on creating and researching on accelerating AI. It is not a group whose mission is to... I think it is interested in safety and I think there is some research going on there, but I think if we come at this field with a mission of saying, we want to help the subfield of AI research that specifically works to reduce risks, we wanna see that field grow in academia, we wanna see there be jobs at all the labs for people who work on that stuff, and we wanna see the field of strategic and policy preparations to reduce misuse. We wanna see those grow and I think while OpenAI is interested in reducing potential risks, that's not their day to day work. And so I think when you look for people who are trying to bring about what we're trying to bring about, and can comparably recommend donations, I think that that's not a crowded field from that perspective. Other questions or did I answer them all? Yep? .

36:08 S5: What sort of policy, sort of preparations are you guys about the AI risks? Is that all like job permeable from AI, or is it about other things as well?

36:16 S1: Sure. So what kind of policy preparations? This is something that we've so far only been pretty speculative about and we haven't made grants yet. We wanna be careful because the thing we don't want to do is start just funding anyone and everyone to start talking about policy around AI. Because for example, I don't think it would be a good idea for AI to be regulated today. I just think it's the stuff that one might want to regulate is too far off, and I think that the harms would outweigh the benefits. So, we wanna be careful, and so, we're very preliminary on that.

36:48 S1: But I would say our primary interest is in minimizing the risk that there's either going to be some kind of arms race type dynamic with AI, or there's going to be a very powerful AI in the hands of basically either a rogue or an authoritarian government, someone with very bad values. Those are the things we're most worried about, it's that dynamic. We... The whole idea of AI causing unemployment, we think is also an area of interest, but we think it's probably going to be a slower developing thing and also is a little bit less neglected than the kind of issue I just talked about. Yep?

37:29 Lewis: Can you characterize how you are allocating your focus between a more speculative Open Phil topics and your more traditional GiveWell work?

37:40 S1: Sure. How do we allocate focus between speculative Open Phil stuff and traditional GiveWell work? So it's... I mean mostly right now it's most staff kind of work on mostly one or the other. I work on Open Phil. So that's... That is my focus, that is my job. I'm still on the board of GiveWell, I still participate in debates about what the top charities should be and how much we should recommend, but my involvement there right now is like pretty high level, and most of my time just goes into trying to build the Open Phil team and work on that stuff whereas Elie mostly works on GiveWell though he also is involved in a lot of the Open Phil discussions, and also is taking that supervisor role on the pandemic preparedness. Sure.

38:30 S5: What sort of discounts do you guys do for human to animal lives, and then also for future lives versus present?

38:36 S1: Yeah. How do we value human versus animal lives, how do we value present versus future lives? Yeah, the more options you put on the table as a philanthropist, the more you confront these really foundational and sort of impossible philosophical debates. I think a general principle we have is to avoid a situation where everything we're doing is only good if you accept a particular philosophical position, in a very contentious debate. I think that is what we don't wanna do and as... Yeah, and as a group recommending a lot of funding, we don't necessarily have to be in a place where we just have to pick one.

39:13 S1: And so, we do try and do some degree of saying "If there are multiple ways to look at an issue, if there are multiple potential reasonable perspectives, if there's several ways of resolving these questions that we could imagine ourselves really believing five years from now, then we want to, in some sense be making some bets on each one." And I think the reason... There's a couple reasons for that. So one is that, I do think that anything you do as a major philanthropist has sharply diminishing returns.

39:43 S1: So, for example, we look at the field of potential risk for advanced AI and we're kind of saying "Anyone who's good who can work to reduce these risks, is someone who we would love to recommend funding to, and we're really looking for people to fund." Kind of similar with Farm Animal Welfare where there's a set of groups and we're

trying to make everything as aggressive and as fast as it can be, but there's limited capacity. And so it's like, if we all decided that we count animals equal to humans and we don't count future people at all, and so we're all gonna work on animals, it's like, I think we could get more done on Farm Animal Welfare but I think we'd hit super diminishing returns.

40:21 S1: And I'm much more attracted to the idea of saying if it turns out that Farm Animal Welfare is the best cause, in some sense it's like something we'll never know, but let's pretend we could some day. If it turns out that that's the place to be or we decide that's the place to be, then we will have already done an enormous amount of good by those lights, we will already have a program. If we decided we wanna put more resources into the program, we'll have the capacity, and the staff, and the knowledge to do that. But then the same also applies to potential risks for advanced AI, and the same also applies to criminal justice reform and pandemic preparedness. I think that's a better situation to be in than to say "Put everything behind one set of philosophical views and say I really hope we're right." Because we think that these views really could change and I think that I... I certainly would hope that anyone with enough resources to do all these things would do them, if their values on these tough philosophical questions were different from mine. And so, that's also part of the thinking.

41:18 S1: So the bottomline is, we do try and take any perspective that we could really imagine ourselves being tempted by or holding in the future, is something that we want to recommend some resources into addressing, and I think that hopefully gets around having to have a totally nailed down view. Honestly my own current best guess and philosophy of mind, and how you wanna weigh animals versus humans, that has varied a lot. But even on days when I don't think that we should worry about animals suffering, I'm glad that we're doing something about it 'cause it would be very easy for me to be wrong. Yep?

41:56 S3: So you talked about if we decide in the future that this is more important change of views in general to the value of information [42:05] ____ value, you might get that information later, any plans to actually support the gathering of that kind of information, or some... Like in five years will there be a big Open Phil funding of research that might prioritize along those lines?

42:25 S1: Do we have plans to fund the research that can help us understand these issues and gain clarity on these issues? Yeah, I would say... This can mean a lot of different things. So one of the ways we learn about our work is, we just look back at things we've done and we see how they went and that does change our view of things. And so, that's something that we do expect to do. It's not where we're focused right now, and it might be a few years before we really put a lot of energy and do a look back. But it's something eventually definitely we want to do, and it's something we do informally now. So, we had a meeting very recently where we all kind of went over one person's recent grants and just informally talked about how they were going. That's one category.

43:05 S1: Another category you might have in mind is... Well, I'll tell you one thing that we're thinking about doing. When you think about estimating, doing kind of a back of the envelope calculation of how much good you accomplished with a grant. There's a whole bunch of sub questions that you run into such as, I won't go into the details on these but what discount rate should you use and how should you value dollars of economic value added to the US versus dollars of economic value added in the developing world. And there's a whole bunch of little empirical questions that can give us more clarity on a lot of how we compare across these different clauses. So, it's possible that after he's done with the criminal justice work that David Roodman will make this maybe his next big project, and he tends to be very thorough, so that would definitely teach us a lot.

43:57 S1: There's a another category of well what about... What are our funding philosophy to try to figure out the answer to philosophy of the mind. I don't know if that's actually possible, I don't know how tractable that is. It is something that has come up, it is not something we have active plans to do. But I would generally say that, if we see opportunities that we think, where we think we can especially if it's more... If we see opportunities where we think investing time and or recommending grants can leave us better informed about how to trade our different causes, we'll be pretty excited. And so, people who know about ways to do that or think about ways to do that should let us know because we're generally open to it. It's not like a super active focused area at the moment. I'll take a couple more questions? Yep?

44:44 S3: My question is about how Open Phil approaches cases where the number of people [44:49] ____ in a cause is relatively small and GiveWell or Open Phil and Good Ventures might have to have an [44:55] ____ power to influence the direction like in the example of [45:00] ____ debate for strategically and morally about incremental welfare approaches like cage-free eggs versus vegetarian vegan advocacy and... So grants of these sizes have significant

influence on all of that [45:15] _____. So how do you all think about that power in some cases?

45:20 S1: Sure. How do we think about power to shape the fields when the field is very small? Mostly we definitely don't want to distort people away from things that they like believe in. We definitely avoid relationships with grantees where we're sort of saying, "Here's something you have no interest in doing and we're gonna make a contract with you that forces you to do it." But something that we are quite willing to do is, if there's a grantee who's willing to go in one of several directions and we are... We have the ability to recommend money that would be a big part of their resources, and we have a strong opinion that we believe is informed about what work is better than other work then we're just... We don't really see any reason why we should do anything other than support the work that looks most impactful to us.

46:12 S1: Of course the important thing there is to try really hard to get it as right as we can, and so, part of that is, I wouldn't wanna be the person with the limited attention I have. I wouldn't wanna be the person trying to figure out all on my own if incremental welfare reforms are more promising than vegan advocacy. But I think Lewis pretty much thinks about this all the time and has heard from the depot on both sides at length, knows the arguments, has engaged in the debates, has the networks, and so, I feel better with him leading the way, me asking a lot of questions.

46:46 S1: We did put out a blog post. On this particular issue which I think is contentious, vegan advocacy versus sort of cage-free type performance, this is an unusually contentious issue for us to wait into and because of that we made bit of an unusual effort to put our thinking out in public. Give people a chance to challenge it, see if there are arguments out there that we are missing and we did have that post on why we actually do believe the incremental reforms look like a better deal than the vegan advocacy. Yep?

47:15 S4: I saw that you recently made a grant to Philip Tetlock to support his work on forecasting. He's fairly well-known, so I was curious why he couldn't get funding from another source and what went into that?

47:30 S1: Sure. Philip Tetlock, why couldn't he get funding from another source, isn't he well-known? It's not that he had no funding. It's not that he was totally out of money, but we are excited about his work. This was a one-off grant. This is not in a focused area. This is just something that, we do a very, very small amount of one-off grants when we have the staff available to do them, and when we're excited about something. And so, Philip Tetlock is someone whose work, a bunch of staff we're really excited about and really familiar with.

48:02 S1: We asked him, "Are you basically funded to do everything you'd like to do, or is there stuff you'd like to do that you don't have funding for?" And he answered the question, he gave us a proposal. We thought the proposal looked like something worth funding, and so, we did it. And that's our general approach to the room for our funding situation. In general I would say it's, I don't think it's usually a safe bet that people who get attention or press are well-funded. I think it's a bit of a complicated relationship. And certain kinds of press or certain kinds of well-known-ness are really good at bringing in the money, and other kinds don't correlate as well as one might expect.

48:40 S1: Cool. Well, I don't see any other questions, and it's 9:06, so I feel like I should let people go. And I will hang around for a little while. I have to run at 9:30, but I will hang around probably 'til then. So, thanks everyone for coming. We really appreciate all the engagement, all the great questions. And people like you who make it possible to do our work are just really awesome. So, thank you.

[applause]