World Cafe ­ Evaluation – Sessions 1, 2 and 3

**Session 1**  
  
[background conversation]

Thank you all for coming up all the stairs to talk about...

I took the lift.

Did you?

[laughter]

[background conversation]

Yeah, so evaluations obviously, and hopefully you've had a chance to have a look at the worksheet here. So yeah, what did you think of the questions? Were they easy? I think we already discussed how they could be interpreted in different ways.

I just want to make sure that I got right, when we mean evaluating SSH into the projects. So how SSH has been integrated in the activities, you mean through [0:01:01] [inaudible] as well right, or so really in terms of topics that have been addressed maybe in the projects or in the policy itself?

Something, yeah...

Because that, yeah, slightly different. Okay.

Yeah. So the particular focus of this session is to think about research that sets out to evaluate something. That's what we mean by evaluation. Any kind of research that is evaluation research. Whether it says that explicitly, or maybe it just says we will monitor the outcomes of whatever, or we will assess the benefits and cost of something that's happened. That's what we consider to be an evaluation.

Okay.

So, then what we're thinking is, okay, given that there are lots of different types of evaluations, projects, programmes, policies. What are social sciences and humanities doing within those evaluations? And that could be different in all those the different types of areas. And are there things being, I mean, is social science having an integrated and mainstream role in those evaluations, or is it actually more hard science subjects that are kind of dominating the process of doing evaluations? And yeah, these sorts of questions then about, how could we do this better, how could we get more social science and humanities work actually integrated into evaluations at all stages. So, the design of the questions, the implementation, the outputs.

May I ask the rationale for why you're asking this specifically, the evaluation?

Well, I think what we were trying to do is dig a bit deeper into how social sciences and humanities are being used?

Okay, in different aspects.

And so obviously there's different types of research going on and being used in different ways, and it seems like evaluation is one particular way. And as it obviously directly often feeds into policies and programmes in the future, it's a very relevant area of research. I say that... I was actually involved in choosing this subject. [chuckle] So that's why I like it, it's interesting. And obviously we've got another session that you'll be going to that's talking about reviewing. Literature reviews and evidence reviews as a sort of another way of doing research. So anyway, that's probably enough of me talking, really. Yeah, so what did you think about the first question, [0:03:03] [inaudible] based on different experiences?

If I may maybe kickstart the discussion because I did myself evaluations of policies and programmes in the past, of EU high level policies, let's say, in the environmental or energy field. And I have to say that in this sense I'm quite surprised how much SSH actually does play a role in my mind during the evaluation phase, as opposed to maybe the phase when the policy was conceived. Because as an evaluation professional, you really have to think of the... Yeah, especially the outputs, outcomes and wider societal impacts of the policy. So then really, you do focus on whoever may be affected by this policy, different societal groups, players, etcetera. So, you really think of them, and you work your way from very specific outputs that is visible immediately, such as learnings, to outcomes that are maybe a bit more longer term, jobs created or something, to wider societal impacts. And, yeah, I think the impacts, there's usually really a focus on society, so what does that mean? So, I think, yeah.

And the funny thing is that when you do this kind of, what is it called, evaluation matrix or something? When you go through the different steps as an evaluator, this is most of the times the first time that the policy makers who came up with the policy getting aware of this. So, they usually haven't thought about this when they conceived and developed the policy. So, this is the first time, usually [chuckle] you're going through with them and say, "Okay, yeah, yeah, has impact on this and that people or group." Yeah, so to my experience, it integrates SSH a lot, but mainly thanks to the evaluators who are usually social scientists. Not usually, but often, or political scientists, or people who work with policy. And I think the methods themselves used during evaluations are also really a mix of quantitative and qualitative methods, such as questionnaires, interviews, focus groups. Of course, usually standard methods, right? I'm not talking about very out of the box methods, but... So, yeah, that's my experience.

So, if these social scientists are coming in just at the end then, are they being told what questions they should be asking or are they actually having some freedom to say, "This is how we're going to evaluate the research, we're gonna judge it by these criteria."

I think there is some kind of freedom. From what I have experienced it's really that the social scientists, the evaluators propose certain things to the policy makers, and they usually accept.

That's interesting because I worked with a couple of people on a study of how evaluations are taken up in policy. And we found that indeed a lot of the evaluations when they are not necessarily done within the policy institutions, they are far broader than just measuring the distance to target that we set out in the policy. So external researchers and think tanks, and consultants, they tend to go far broader than what does this policy aim to do. But it doesn't necessarily get taken up in the next round of policy and that's what we found. So, it happens.

It's not necessarily prompted by the policy makers themselves and when it's communicated back to the policy makers, it doesn't necessarily influence how they design the next round of policy. And that was a study that looked at environmental policy evaluation, so potentially specific. And it looked at EU environmental policy evaluations over certain period of times. So, there's always these... This was the general conclusion that we come to when we looked at the work across.

So, the actual evaluation happens indeed but when it's outside the policy's sphere or when it's outside the policy institutions, especially for example this REFIT thing, which is really dominating having indicators and quantitative measurements for everything, although they use qualitative methods a lot of the time, still they need to have very specific data about it. But if the evaluation happens outside of the policy's sphere, we find a lot of broader aspects, so societal impacts, as opposed to just distance to target but that we find with the learning aspect which goes back into the policy making world is where it falls apart.

And that would be... Something that, trying to put together a research grant proposal on something like that how to make policy evaluations more integrated into the very agenda setting apart of the policy cycle which we haven't quite figure out how to do yet. I don't know if it's...

That's really interesting, we might need to talk to you more about that, 'cause that's obviously really...

[overlapping conversation]

Isn't this what the aim is for better regulations etcetera so that in the end also... That's the idea behind it I think. But also, for the next impact assessment, let's say of the next policy cycle... In theory, it's all they should use the evaluation outcomes but to what extent then it's really not?

The other example is the formal programme itself for Horizon Europe they've gone through several processes or evaluation of Horizon 2020. [0:09:31] [inaudible] the whole [0:09:32] [inaudible] approach that was brought by [0:09:36] [inaudible] all the research that she made on missions and how we can better address a policy towards a mission that is clearly defined and is interdisciplinary, has a social... Strong impact from social point of view. I think it is a very good example of how you can learn from what you do and then try to then make new research at the same time and then shape the new policy around the results that you had.

There have been quite many changes in the Horizon Europe programme that have been really inspired by this evaluation of Horizon 2020. Not so many could be implemented because in the end they decided to go for a more evolution than revolution of the programme, because there are so many people that are used to already have management procedures that are associated to Horizon 2020. So, you can't change everything, of course, but I think they've done a good job in taking these results and also developing farther this integration of social sciences into the Horizon Europe framework.

Okay, and who did that evaluation or?

So, I think was [0:10:52] [inaudible]... Yes, it was driven by the commission, but I think there were different consultation process for instance... Here I'm thinking now about the work we are doing in [removed for anonymity purposes], I wasn't here in [removed for anonymity purposes] last year, there was a broad consultation with the members of the organization that then fed into the Horizon 2020 assessment process. So, it was done by the commission, but it was also open to public consultation as well. So other organisations they could fit and say what were the main issues with the programme and this could be tackled in say the next one.

That's interesting, so kind of balance of internal and external, maybe making sure that it did get listened to because it's got that internal aspect to it but also there's the voices...

I was working mainly on the simplification, for instance, of the programme so it was really, I would say from a management point of view perspective, I wasn't really into the social impact and other topics, but I could say that the commission to come mainly of the recommendation we made, then into the new regulation. So, it's good, also you see that... [laughter] the feedback is listened to.

But again, that demonstrates that, yes evaluations they're usually really open to the SSH input that exists, such as public consultations etcetera.

Okay, yeah.

Again, as opposed to when they conceive the programme because then usually it's just a political decision and... They're not even clear on exactly what would be the effects, the impacts.

Okay, and do you think that's all the different SSH disciplines, then would be able to feed in at the end?

I don't know from a methodological point of view it's more probably social science, such as... I mean using methodologies that... Questionnaires and qualitative semi structured interviews [0:13:06] [inaudible] I don't know if you want them... From this point of view, it could be more like the standard social sciences and maybe less humanities, or...

Yeah.

Yeah. I think also then the impact that you're looking at are probably also same in terms of how the policy was worded. So, if they speak about consumers...

Mmm­hmm.

If you're not certain... Suddenly starting to look at how values change and... Yeah. How about you, guys? Have you been involved in doing any evaluation kind of research?

No. Not yet. So, I don't know anything about doing evaluation research, but I thought about it in the political process. And I've been wondering, for example, I've worked on the issue of social acceptance. Social acceptance of wind energy, for example, and there is a new law, or an absolutely new approach tried in one federal state in Germany. And they... During coming up with this law, they put in a sentence to evaluate the effect of the law after a period of three years.

Right.

But it was a parliamentary evaluation, so it would...

What is that?

Political evaluation. So, it has to be discussed in parliament. And I was just... Because of that I was thinking, and of course they have funding and they... Because we were kind of involved, our institute, in coming up with the policy. They asked us if we would be interested to learn if a call would be issued to evaluate the law. And I was thinking, is there a difference between commissioned evaluations and independent evaluations which probably have been driven by purely academic research interest, for example, political science books. I guess, the classical question is looking at the impact and effect of policies and how these two, let's say, starting point for evaluating a policy would compare? Just as an open question.

Yeah. Yeah. I'm sure they would be very different. I mean...

Yeah.

For a start, which questions actually can be asked, would be very different, [0:15:42] [inaudible] whether it’s commissioned or it's independent.

Yeah.

Interesting. Yeah.

I think it's more and more of [0:15:46] [inaudible] to have these evaluations done by independent...

But actually, we looked through the... We also... When we look at... So, when we were talking about environmental policy, [0:15:57] [inaudible]. When you look at the environmental policy evaluation, the amount of policies that are actually evaluated is a tiny proportion of the overall policy. From an evaluated ex­post so after implementation. All of the... At the [0:16:14] [inaudible], they're all expected to have impact assessments. There always is some sort of monitoring and reporting, whether a review evaluation from the perspective of, thinking about the impacts beyond the policy is still quite a limited percentage or proportion of the overall. At member state level the proportion's even lower.

Very few member states systematically evaluate. So if there is any type of evaluation written into the policy, that is a requirement, that's already a step forward.

That's what I thought too. 'Cause I haven't seen that many [0:16:44] [inaudible].

It doesn't happen, often. And if you speak... It depends on the state. I think the Netherlands, the UK, [0:16:49] [inaudible] it depends on the land. For France, there are some state, members states who were better at systematically evaluating the policies than others. And if you speak to the civil servants who are doing it, they're like "Well I don't want to evaluate this policy, in case I find something that upsets my colleague."

[laughter]

It's only a very... It's a silly anecdote. These things become barriers. So, I think it's quite interesting that that's already in the policy. That's already a step forward in terms of taking the evaluation seriously.

Maybe... Also adding that... What you just mentioned France's Horizon 2020. So that was the midterm evaluation, and... Because it's obviously not finished yet. And most programmes or initiatives, they get evaluated just after they finish. So, you're just guessing what's the longer-term societal impact will be. But it's usually just a guess work. And it's true that there are very rarely any evaluations that would go... That would be done five years after it ended and then revaluate the impacts which would...

A lot of the time, the sort of monitoring and reporting aspects of the policy is only over the lifetime of the policy. So, if you want to evaluate five years after the fact, perhaps there will be data gaps that you can't rely on. People have moved on.

So, it's quite a difficult exercise to do on any of this.

I think for academics, as well, maybe there's a sense that a policy from five years ago, what... Where's the kind of excitement in evaluating that? Is that just me, but it's old news, isn't it?

I think in the academic world, at least you have the possibility to do that. If you're in a... Some legit consultancy agency or research think­tank sort of world, it's even less attractive. At least the academic world, you should be able to have the need to investigate things in group from a distant past if necessary.

I was interested in what you were saying about the piece of research you did on the take up different evaluations?

So, I wasn't a research that I did... It was research I was involved in one way or another. There is a PhD research... He's completed his PhD research. He's done a lot of this meta-analysis of evaluations and evaluations. He's based in [0:19:12] [inaudible] University. So, he is now the expert in that field. So, the research that I'm talking about is his baby that I was interested in following.

Yeah.

I can give you his name, if you want. He did his PhD at some university [0:19:32] [inaudible]. Ah, really? Okay.

He is [0:19:37] [inaudible]. Oh no, it's Schumacher, because it's German, not Dutch. Sorry. Ah, great.

Can't wait, what his...

Has he finished the PhD now?

He's finished the PhD. He's now working in Germany.

Ah okay. Yeah.

But him and a group of other people around him and us... There's a network of environmental evaluators. European Environmental Evaluators Network, which is a network that comes together once a year and all talk about evaluation, how, like progress in the field of research. There's an annual conference and we all work together on these things, and we raise pretty much the same questions every year. All of which is about systematising evaluations across policy fields, implementing it into the agenda setting of next policy cycles, uptake of the findings, replication, data gaps, the same sort of questions come up every year.

So, this is a whole group of researchers across Europe. I think there's an even international... An international sort of arm of this type of network. If you look up EEEN Forum, it's terrible, E­E­E­N Forum you'll find a lot of information. So, the European Environmental Evaluators Network, and that's why I said I'm only talking about environmental policy [0:21:07] [inaudible] when it comes to evaluation, it's the only field I know [0:21:10] [inaudible], maybe it's completely different if you’re talking about migration.

[0:21:17] [inaudible].

Yeah.

That's interesting, I think because those points that you just listed are perhaps some of the answers to question two on the worksheet, aren't they, what do you need to think about to do a really good evaluation, so it would be to do with things like [0:21:31] [inaudible], agenda­setting, ensuring the uptake. Well, I don't necessarily know how those things would be achieved, but that is sorts of things that it would be needed to address. Did anyone have any examples of... We've already discussed the Horizon 2020 into Horizon Europe as an example of a good evaluation process. Yeah, were there any other ones?

That one is nice, it's kind of setting the standard, I'm trying to think about...

[chuckle]

I'd like to learn a bit more about that one actually, you said that it was an in­house one, but with consultation.

I think so, I'm not sure whether there were also some other external reviewers. I don't know.

[overlapping conversation]

Although it was really big, like this midterm review was something really that big and advertised a lot so...

Okay, yeah.

For sure there were consultation processes, so everyone [0:22:36] [inaudible] comparing and say [0:22:38] [inaudible] what were... But I'm not quite sure whether it was only done internally within the commission as well, or if there were some other external...

[overlapping conversation]

I know that my university contributed to the consultations. I don't know if it was spontaneously presented or if it was organized, but I knew there was contributions.

Plus, there were some other consultation policy maybe not only online organized, so anyway, the commission was getting in touch with the stakeholders' organizations, so people, organizations, they were highly involved into the thing, programme, in meetings and other, they really had several ways to get views on it. Not only stakeholder consultation maybe online but also events and so on. So, this was a huge [0:23:19] [inaudible] consultation process.

Yeah. And in terms of the social science and humanities role in that, was it very integrated then? Or was there a sort of a bit of work set aside that was...

[0:23:31] [inaudible] to what extent, let's say, if we look at our topic of today, to what extent have they evaluated? How, for example, SSH was mainstreamed in the results. So, I mean we know now there’s isn't methodology exists but... So that [0:23:53] [inaudible]...

There is a question to take then with [removed for anonymity purposes] will be [0:23:55] [inaudible] to ask to him, because it will show, you will definitely know how much SSH has been, the project officer will know how much SSH has been involved in that, I don't have a specific answer on that.

[overlapping conversation]

They've been looking into definitely gender balance, geographic aspects for sure. I'm pretty sure there must be some quantitative data also and SSH involvement, but maybe he would have a good answer to that.

[background conversation]

So, I think it is in fact time to move on to the next cafe. But is there just any kind of final thoughts? Do you feel like we've been covering the right sorts of questions in this session? Because obviously when new people come in, I can ask different questions. So, if you think...

I think it's interesting that the title of the cafe is evaluating successes and failures and we've been talking about evaluating impacts which is maybe a bit different because I think evaluating successes and failures implies that we're evaluating how successful a policy, or a project has been in meeting the goals that it set up for itself. What we're saying that the SSH evaluators do is assess what the impact of this project or policy has been on wider society. It's far more than this idea of, "is it a success, is it a failure?" That's, I think, a different perspective, which is what the SSH brings that maybe beyond cost benefit analysis which economics is [0:25:46] [inaudible] social science too but the broader perspective.

Adding on that, I was feeling in the beginning, not in terms of evaluation because I didn't get probably the question right, but I was thinking about the set plan and, in general, energy technology projects. And if I think about different implementation plans that have been developed so far, we don't have evaluation yet, but I can say that social sciences and humanities sometimes are involved as cost cutting activities. The main objective indeed where the success will be measured is technology driven, so that's exactly what you were mentioning that we have to make sure that learning the evaluation, I think the impacts on society are also evaluated in terms of... Because, if we think only about the success of the project, of the policy itself, probably they will not be taking them much into consideration if the project is about developing a specific technology in this field. So yeah, that's definitely...

[0:26:46] [inaudible] policymakers, because they have probably this intention, you know the... It's very hard to get over... Technology, and then they are surprised, "Oh there are wider benefits."

[laughter]

Especially when the research is about low technology levels. When it's, for instance, also in here we focus on technology readiness level that is up to five or probably seven which means that these are technologies that are not deployed on the market yet, and therefore there the social impact is not that much... Of course, you take into account, but then when you are at higher level of technology readiness, of course you thinking also about how you put these technologies in the market, how then the citizens will take up the technology, etcetera. So, it's good to have them implemented also then in the first stages, so that you have a co­creation process and it's not only then, okay, what is the impact of the technology now. It brings social sciences at the beginning of the research process, then you could co­create it with the citizens and consumers and that's a different thing. Then the take up of the technology should be more strong of course.

Thank you all very, very much. We need to hurry up to your next group. And if you see [0:28:02] [inaudible], [removed for anonymity purposes]

**Session 2**

[background conversation]

I'm just writing down some of the things the last group said while I remember them 'cause I forgot that I have this [0:31:27] [inaudible]...

[background conversation]

Cool, hi everyone. So, this was just a few of the points that the last group raised, so I just thought I'd give you a very quick recap. So, we were talking quite a lot about when evaluations of projects and policies are done internally by whoever kind of designed and ran the programme compared to when they're done externally, and also the difference then between if it's commissioned by the people who did the programme or if it's actually completely independent. So as in some academics maybe come along and just decide to evaluate something so they can write papers about it.

And the idea that obviously, you get more kind of challenging, wide ranging sorts of questions being asked if it is external, but also that it's less likely to get listened to if it's coming from those people, and sort of addressing different questions perhaps to the questions that the policymakers and project managers were actually thinking about. So, they're more likely to take it up and use it in their future policy development, if it's an internal evaluation.

Let's talk about when evaluation happens. So actually, is it midterm, or is it the end, or is it actually after the end? [0:34:04] [inaudible] long-term impacts and that's getting less attention. And then also, within that, where is the SSH component coming? And not often... Well, one idea was that maybe SSH isn't getting into the starts of projects and the agenda setting and research design. But when you do come to the evaluation, there's this idea, "Oh let's assess the social impacts of our project." And especially if you could show wider benefits, so they might... Their main indicator might be saving money or the development of technology, but then if they can show, "Oh by the way, we've also addressed gender or multiple benefits in health and things like that, they might like that.

And also, just to note that... I suppose this kind of pervades everything we've been talking about with the way that the initial framings and conceptualizations of what people are and what energy systems are, how that kind of follows through all the way into the evaluations. So, the way the questions and the methods are framed will be very much determined by those ideas that underpin the whole project, especially if it's an internal one, obviously. And then just in terms of good practice, we didn't get a lot of discussion of sort of what that is and specific examples, but we did think of one, or at least the previous group did think that this was quite good one is that there was a mid­term review of the Horizon 2020 package, especially trying to inform the development of the new Horizon Europe, and that was seen as quite a wide ranging and quite an in depth review of the Horizon 2020 work that had gone on.

So, I don't actually know about that review. So, you may feel differently if you know about it yourselves. Yeah, so that was some of the things that we've already covered. So, yeah, [chuckle] I think I'll probably do the same thing again and just like scribble notes down in my notebook rather than trying to stand up there the whole time. So, how did you find the questions to start with? Did the questions all sort of make sense to you?

[chuckle]

[pause]

So, we had... Yeah, go on.

I have a question about the first question.

[laughter]

Because I understand what you discussed, but now the way I... I have an issue because like for example, evaluation itself is an SSH exercise in my opinion. It's like a social... It's like we ascribe the value to some effective policies and then we measure how that is done. So, I think that's a certain... At least it uses social science research methods to see if a policy has worked or not. So that's the first point. But then I would like... My question would be, do you mean if SSH evidence is considered as an evaluation, or if SSH on research on a specific policy should be part of a normal evaluation? So...

Yeah...

Or because for projects, the use of SSH outputs might be very interesting, but for policy, that's a different thing. Like for the renewable energy, their record is not gonna count, SSH outputs, I mean, but that socioeconomic aspect. Did that make sense, what I'm...

Yeah, I think I see what you're... Yeah so, we had a bit of discussion actually, yeah, in the previous group and even before we had these sessions about what the question actually means and what do we mean by evaluation. And I think... So, what I mean by evaluation is research or data gathering activities that set out to assess the impacts of a particular policy or project. And so, what we're interested in particularly in these sessions, although this is kinda prescriptive, but I think the thinking behind it is when those research activities, as evaluation research activities happen, are they being done with social science and humanities integrated into that evaluation process or are they not. And I thought your statement was really interesting, all evaluation is social science and humanities. You have basically suggested with that that is integrated, and we can't... Yeah.

Well, I think shouldn't... I mean it's a social science research method. It's like measuring certain things and then interpreting it in some ways, in my opinion, it is.

Yeah, go on.

It's not like a scientific research question of course, but it uses methods from the social sciences. Interesting. Do you wanna...

Yeah.

[overlapping conversation]

I that that's really interesting. I also... I had problems with the whole AE thing 'cause this evaluations thing is not really familiar to me, but I tackled it more from a quantitative versus qualitative perspective...

Yeah, yeah.

Which is as [0:38:54] [inaudible] said is not truly accurate because definitely there is also a quantitative social science, but I do more quantitative social science, that's why I tackled it more from this perspective. And I try to [0:39:13] [inaudible] but what would be benefits of integrating more qualitative evaluations in a more... Instead of only quantitatively assessing or evaluating certain impacts of... I think one of the benefits for me would be that there... It's possible to have a more... A focus on more complex explanations, instead of kind of strict causality ideas which might be... Because we already had this in the group before, because one of the things that social scientists always say is, "Well, it's complex."

[laughter]

Yeah, but in evaluations, this might also be important because very often it's not only one factor explaining why a certain policy or project or whatever succeeded or failed or had this or that impact, but we could have more complex view...

Yeah. So, yeah, so I think this is really interesting, like what actually is an evaluation and what is the role of social science in that kind of... You know, inherently in that approach to research, and I think maybe if the term "evaluation" and the term "evaluation research" sort of implies a specific kind of body of methodologies, which really do involve a certain degree of social science and humanities, but I think if you look at the sort of policy worlds, when you talk about impact assessment or measuring outcomes, those are essentially evaluative activities and they are very much not inherently social science, I would say. A lot of the actions that are in the set plan, a lot of the current things that the ETIPs are doing are about "right, we are going to get from here to there in terms of technology or in terms of patents and markets."

Yeah.

And then obviously it depends whether you see economics as a social science or not, but there's a lot of evaluations I think that are basically just pure economic evaluations or...

Yeah...

Yeah, what do others think about that?   
  
For me, economics is a social science...

Yeah, okay.

[laughter]

Fair enough yeah, okay.

I see, it depends on the policy. For example, I think that for example the evaluation of Horizon Europe is different from the evaluation of the renewable energy directive, or from something that evaluates the impact of whatever, land degradation, that's very different, so what the subject is of course. That of course defines the tools for the use of evaluation as well...

Yeah, okay. So, do you think that social sciences and humanities have a different role to play than in those different evaluations?

That's a good question.

Well, [0:42:15] [inaudible].

If I understand correctly what I've been hearing the last few weeks, because [0:42:23] [inaudible] who is the director general of the GITD and who is, nowadays running from one meeting to another to present the process and the procedure about the planning of Horizon Europe, what he keeps saying is that indeed in H2020, it was all about those SMART submissions would have... The proposals would always have to put SMART, like targeted impacts presented at SMART so it's measurable, it's something that...

Ahh, okay, yeah.

S is for I don't know what anymore, M was measurable...

Specific?

Specific, measurable, achievable, blah blah and...

Time specific.

Timing! Timing, something like that.

Don't know what R is.

Reliable or... I'm just making up...

[laughter]

Whatever an R can be...

Yeah, yeah.

It was something about like... It was a...

Replicable.

We will take up that amount from that percentage to that percentage increase in, and that was really specific and measurable. And what he said was that from the valuation of Horizon 2020 which is not yet finished, but still... That doesn't work because even though everyone agrees that this is, has been and is still a very good framework programme for research and innovation, in terms of those impacts, almost none of them have been achieved. If you have to say we will achieve that increase of 15% in take up, I don't know what. Everyone is... That's a bit what I was saying in the other group, nowadays researchers but also innovators, they need to oversell their presentation.

Yeah.

And so, everyone knows that what is written in the research proposals, it's not doable, but it's not doable it's in the three years of time or framework of time, or the 50... Five years. Everyone knows that, even the people who write it down, the people who submit the proposals. So, everyone knows that, but everyone keeps doing it because it's the way you have to...

Right.

Do that.

Yeah.

And so, then when it comes to evaluate the impact of something that everyone knows that it won't be achievable. Of course, it's not achieved in the end, but it was not supposed to be. But yet it was written as that.

Yeah.

Apparently, they've been aware, they tend to become aware of there is something problematic in there and it should not be in the same way for Horizon Europe. How it will be, I don't know, but they have... They [0:45:41] [inaudible] that there is something not going well with the way that... I'm not sure they would say the way I'm seeing, it's that everyone keeps overselling everything...

Yeah, yeah.

Because that would... I'm not sure that it would be politically correct to say... [laughter]

It is what it is.

Yeah, yeah.

And I don't know if that is useful in any way...

Well, yeah because it's about how you frame the aims and the success, kind of criteria of programmes is inherently bound up with how you then evaluate them because you... Especially if you're a... Well, internal or external, you've got to evaluate against some kind of aim, haven't you? Yeah, if those aims are nonsense then where does evaluation then go? Yeah, and in some ways, maybe that gap does leave room for social science and humanities to look at... "Okay, so we're not gonna worry about whether you hit that target, but what else did you do?"

Yeah.

"What other impacts?" I don't know, what do you think...

As I said in the other working group earlier, also evaluation, there is a lot of literature that sets out the best practices and what to consider and so on, so on. This is certainly something that you can already tap into, and there are also evaluation standards, at least in the German­speaking world, I'm sure it's the same in the English Speaking world, there are specific associations or societies or whatever that basically do nothing else but for research institutes that do evaluations, prescribe what are the standards that you have to keep, to uphold in order to make a sound evaluation of whatever and of course, a very important part of that is always that you have terms of references, terms of references against which you evaluate something.

So the terms of references usually are then also negotiated between those who do the evaluation and those who are evaluated or ask for the evaluation. And inherently, many types of evaluations are usually very political and the framework programme itself is the best example of it. Whether it is the midterm review or the final review or whether it is one of the 300 specific evaluations of specific programmes that are going on during each cycle of each framework programme edition, it is always that the outcome is somewhat critical to some aspects and to somewhat it is always in total, this is a great achievement or so.

This is, we have... I'm living now through... If I capture my life in cycles of framework programmes, this is the third cycle I'm going through, and I'm witnessing the same things that I've seen for the FP7 and for the Horizon 2020. It's coming every seven years again, two or three years before the new framework programme starts. They usually have this somewhat self-critical and then upbeat approach, next time we will make everything different, the next time it will be conceptualized. This and there are certain buzzwords around that are then forgotten for the next time, last time it was RRI, for example, very important, no one is talking about RRI this time for [0:49:05] [inaudible] Europe. But impact of course, is very strong. For FP7, one buzzword for example, was excellence. Now excellence is included already with the ERC, so this case is closed, but that just gives you an idea of how these... And the framework programme of course, is huge. It's 70 billion, or next time it's 100 billion or 90 billion, whatever, but the same logic in a way applies to different programmes of different scales and different whatever you evaluate.

There is always something inherent political in most of the evaluations. I would not agree that the evaluation is an SSH exercise, I would agree that a lot of those who do evaluations professionally are people who have a trained SSH background but evaluation per say, as far as I can see it, in the history of science is basically something that has been done in order to look at what worked and what did not work. Or does it work or does not work? And this has been done in the past by hard scientists as much as it has been done by SSH people.

Just picking up on what you were saying about these different cycles of the framework programme. How do you feel about the role of social science in those kinds of exercises, those kind of self-critical, self-reflective exercises? Do you think it's playing a role in that?

Well, I should say upfront that I am not too happy with the label SSH, because this is an administrative creation by the commission, and it brings together very different strands of research and epistemological and oncological ideas that fit very uneasily in my understanding. And if you ask me, and I've done a little bit of historic research on the framework programme when I wrote a book on one specific programme there, what I saw in the record, but I did not look at this systematically, but what I saw, those who are mostly involved in prescribing what the next framework programme should achieve and giving kind of the ideational framework for the debate are usually economists, for a great degree.

This of course relates it very strongly with the overall idea of the framework programme belonging to a certain budget line in the EU budget, which is called competitiveness, so this is a fact. It is called competitiveness and within this of course, you have certain strings attached already. This is the meta­political area. And the other people that I have seen quite often are sociologists, political scientists and STS people, so science and technology studies, who then are more not on the, maybe not on the ideational level, but they are very often working in specific areas to argue for all these other buzzwords that are out there. What was the best­seller of [0:52:26] [inaudible], a novel in the '90s with Stevens. [0:52:28] [inaudible] or triple helix, quadruple helix, all these things that kind of given... Are intended to make... To give sense, to give purpose to these kinds of exercises.

Okay. Triple helix?

Triple helix.

What's that?

Research, education, innovation...

The idea is that you have... Now it's a quadruple helix, it means that, you have interplay. Universities, governments, or industry and public and society at large and be together...

[0:53:09] [inaudible].

Prescribe... Describe the agenda of what research [0:53:13] [inaudible].

Okay, yeah.

But I mean, really these are the... If you look at the disciplines, it's really... And also, the evaluators of the framework programme, most of the time they were high level natural scientists and usually economists.

Right, yeah.

[0:53:34] [inaudible]. Me and [0:53:40] [inaudible] 'cause now everyone...

But she did not do a review. She wrote one of these white papers.

She wrote about the innovations on the missions because she's an economist so...

Of course, in economics, you also have different shades. She is of course, much more post Keynesian or Keynesian whereas in the early 2000s those who were more in charge of this were more what you would call nowadays the mainstream economists would more look at... On the consumer side, where the consumer thing again.

It's kind of power shifts and cultural shifts then...

Well, the entire framework problem is the negotiation of different interests and that's one of the reasons why it is so... In political terms, why it is so successful in my understanding because it is not a single programme, but it is, in fact, a huge box of different interests that are related and somehow also served with certain amounts of money and those different interests are negotiated three, four years before the next framework programme starts. That's why all the lobbying already starts for three, four years before the framework programme starts. And it is great because they can give all these different communities their resources that they're asking for. And yeah, sometimes new initiatives come up and try to speak on something new but overall, it's in political terms, it's one of the most successful. For researchers is not often that...

Very often not very satisfying because it has all these strings attached and all these... Is bringing all these difficulties with... Almost no one I know has a complete oversight of what the framework programme actually contains. I just looked by chance on the CORDIS website. You know CORDIS which is the database with all the projects and that you can also click on programme, and I think I ended up with over 300 programmes that are currently running under the Horizon 2020. So, I mean, in that terms it's very frustrating because you belong to one of the niches that are served or you are completely outsider.

Yeah.

But there was not too much about evaluation or... Sorry.

No, no, no. It's all relevant and I think you had the question of what even is evaluation is quite an open one isn't it? So, I think it's probably time to move on to the next one.

Sorry.

No, no, no unless anyone has just a final few words on kind of good practice or any examples of good practice to share. Sorry, we've deconstructed the question, haven't we? Who knows what good practice is?

[laughter]

No? That's cool. Alright, well thank you very much and...

Maybe I can say one thing. This is very much because I was working for the ERC beforehand and that's why I'm biased there. But I think one of the lessons that you can learn from the ERC when you talk about evaluation process of publications, which is really very, I think, a very sound measure is the fact that if you evaluate a programme or a proposal even from a certain specific area that you then have evaluators who have a deep understanding of that field. So, in terms of the ERC...The ERC does... Funds research from all of the different disciplines, academic disciplines, but it is always... If you have a proposal in political science, then there will be political scientists and sociologists who do the evaluation and not chemists and physicians.

Which can be [0:57:41] [inaudible] itself.

Yeah. I mean it's still better than if you have someone from life science and he says or she says "Well, the H­Index of this person... " What I'm trying to say is that these people have... They know to measure... Their peers.

Their peers.

Just their peer... Because they know what is the reputation... What is the currency in this field.

Yeah, but the thing is that, if you have reviewers and natural sciences that are from the say, physicist, from physics, I think it's clearer than if you have political scientists or sociologists for sociology or political science.

Of course, because it seems that much more fragmented. Yeah.

Yeah, much more... And that might be a problem that you have the same peer review system for these different disciplines, or you have a lot of kind of conflict between different theoretical perspectives that you don't have or not in this way in chemistry or [0:58:48] [inaudible] maybe...

[chuckle]

Then let me rephrase more abstract, it is always good to have... To have a... To make sure that whatever is evaluated is evaluated by its own standards and not by outside standards like abstract H Index or whatever. 'Cause this is something that's happening. This is happening quite often.

Well, thank you all very much. Hopefully you all know roughly where you're going. It'll be downstairs somewhere.

Downstairs. Well, this is the third and final one so we should know where to go by now. Yes, it's the one you haven't been to yet, yeah.

[laughter]

Great and tell people to come up here if they're supposed to be up here.

[laughter]

Right so it is down.

[background conversation]

[foreign language]

[pause]

**Session 3**

[background conversation]

Okay. So, we've had two conversations about evaluations already of which, actually quite a lot was discussing these questions I've written down here. What is an evaluation? Are all evaluations social science and humanities? That was in the last session, we had a discussion about that. And what is social science and humanities? Anyway. So, I think "What is an evaluation?" is one that everyone has come up with their own different sort of understanding of what we're actually talking about here, so that is what you wanted...

[background conversation]

[laughter]

So, the big themes that have come up so far, I think this is probably the big one just for that how political evaluations are, and how much they are determined by the agendas that were actually driving whatever policy and programme is being evaluated in the first place, especially when it's being done internally in­house or it's been commissioned, paying somebody to evaluate you. With this kind of rare beast of the independent evaluation perhaps having a different role, perhaps being more wide ranging, more potentially critical, but also those internal ones are much more likely to be feeding back into the future policy and programme development. Possibly just because they are speaking the same language and working with the same kinda frameworks. Yes, so we had a bit of a discussion about when evaluations might happen and when social science and humanities might come in with the idea being that sometimes you get... Maybe the agenda is being set, obviously by the policy makers themselves or by certain kinds of economic type sort of disciplines but then also at the end of a project, that's often when they do open it up a bit more, and say "Oh what are the benefits... In particular what are the benefits might we have worked out through our project?"

And we've had bit of discussion about good practice including some examples, there was one idea that the midterm review of Horizon 2020. Which has been feeding into Horizon Europe, that's seen as perhaps an example of good practice. That being an in­house evaluation, apparently but with experts feeding in. Thoughts about that? And then just a bit of discussion recently about whether the evaluators should be basically from the same community, from the... Very closely connected with the thing that they're evaluating, which obviously avoids the problem of social science work being evaluated by natural scientists who just have no idea what they're talking about. But then also perhaps a slight concern that you might get this series of silos then, with nobody actually challenging each other. Those were a few of the things that have come up, and obviously, the link between the evaluation at the end, if it is happening at the end, but with the very early stages of designing the project or programme, that if it doesn't have good goals and targets built in, then any evaluation is gonna be problematic.

Right, so it's just the three of us. So don't think you can sit back and just you know... [laughter] No free riding on this particular discussion. No sorry.

[laughter]

I read quite a lot of on this one 'cause it was the first one that I did so...

[1:10:47] [inaudible].

Excellent, were you thinking on the same sorts of lines, do you think, as what the people have been saying or have you interpreted it in a very different way?

For the first question, I was interviewing a guy who wrote the book "The Power of Science," who's still working for the commission.

Okay.

And he was going SSH energy issues at DG RTD and back then it was not SSH, it was socioeconomic.

Uh­huh.

And more economic and he wrote this whole book about how socioeconomic evidence informed policy making, energy policy making.

Oh right.

And when I interviewed him, he said, I was asking about more qualitative aspects and he said it's all about foresight in modelling exercises, which have some I don't know of qualitative ideas and not just numbers. So, this is what I know whereas a book about how this evidence play into real policy making of energy. But this is the kind of social science, but they consider having evidence for that.

This is very narrow.

It's really narrow, but it's for I don't know for how many this is the case, so it's very hard to change this mindset into coming up. And even within the social science community for example in sociology, there is a huge group who says that there is really hardcore qualitative research is not as valuable as some other statistical evidence. We already have it in the discipline then between disciplines and then within the policy context, so I think it's a really, really long-term thing to being able to integrate some more qualitative methods I think is the key.

[1:12:55] [inaudible] evaluations, so...

I mean the qualitative/quantitative thing came up in our last discussion as well, and so somebody sort of apologized for saying, "Oh, obviously I'm being too generalistic, but I think we need more qualitative evaluations [1:13:07] [inaudible]." And it's interesting isn't it, because you could probably say that almost all standard science evidence is basically quantitative, isn't it? But with social science evidence obviously it can be either. So is there maybe a difference, and with the quantitative aspect of social science and humanities, getting a lot more uptake, and lot more use...

[overlapping conversation]

[1:13:29] [inaudible].

Yeah, economics. Yeah, so a very kind of differentiated picture really...

Also sociology, they have a huge part, and political science, they have a very [1:13:39] [inaudible] studies about [1:13:40] [inaudible] like 25 countries, how do they use energy, and then you have all the numbers and how it changed over the last 20 years. There's a lot, and yeah, that you can have one figure and it explains everything, you need to read the story.

Yeah, yeah. And that was basically what this author of the book was saying that it's all kind of... You have to have... You have to speak this language essentially of modelling if you need to be listened to.

Yeah, I mean he was maybe at this point not so critical about it. He said it's like a good start to at least have these aspects within the evaluations, that it's not purely numbers.

Oh, I see, right. [chuckle]

So, it's like...

Yeah.

It's not a team that's very critical about it.

Yeah, yeah. And [removed for anonymity purposes], have you done evaluations? Or...

Yeah, I've done some evaluations and I probably got my own sense of it, but I also probably, on the same question, probably try and open this up quite a bit, but let's see. But some of the stuff from the other groups does relate a little bit to what... Well, the first question, in your experience, how does SSH play into evaluations? I mean, similar to what you're saying and kind of previous messages, it's often brought in in this kind of service role, it often doesn't set the agenda or frame the terms of the evaluation. You can see that in the previous group’s comments about the agenda, bla­di­bla. As we hear here naturalistic social science dominates, from my understanding, and the SSH contribution. So SSH, beyond naturalistic social science and... I kept saying that throughout the day because I mean... By naturalistic social science I mean social science that tries to mimic the natural sciences.

Yeah, yeah...

And for me, the qual­quant distinction is not as important as the move to interpretive modes of enquiry or constructivist modes of enquiry in these kinds of things where we move into an area of social science humanities where we understand that the way in which we know the world is constructed by the ways in which we try and access it. So you can be so called reflexive about qual and quant. It's not... And qual can also massively kind of close down and do the sorts of things that might not like it to do, for instance.

So, most evaluation kind of notions come then from natural physical sciences and naturalistic social sciences which has this kind of linear cause/effect kind of model of what you're trying to evaluate. So, if it's a policy, you do the policy and then it leads to change. If it's about, I don't know, behaviour change or any kind of engaging society with energy, it's like you do something, you provide information, it leads to a knowledge gain, you get involved, people involved in a behaviour change programme, it leads to them switching lights off, you're cutting carbon, you get involved in the deliberative process and that impacts on the policy that's being made, or not the case may be.

So there is this kind of linear, and therefore, there's very narrow kind of pre­given understandings of what success might mean in the evaluation that's at hand. Now, beyond naturalistic social science, fundamentally, problematises that linearity in terms of cause and effect and sees things that they can multiply productive. So, you get this where you say towards the end of projects or policymakers like the other benefit. The thing is, why is it surprising that any of these things of engaging society could be multiply productive?

Someone could engage the consumer, and then start getting activist [1:18:01] [inaudible] then because something might... Do you know what I mean? Anyway. I won't go on any further on that. But that was my response to the first question in terms of the state of play. And the other point then with that is that I think evaluation then works often within this decisionistic kind of framing and natural physical science and naturalistic social science that does this as well, which is to say that what...

And also, this is like when we're giving evidence to policy, it's about, we're supporting particular decisions. Whether this is policy decisions, or decisions made by individuals in their own behaviours in their own daily lives, they're still discrete moments of decision and what social science humanities in the broader sense, pushing through into the humanities does is, this is a fundamental problem, because it's these things that we're looking at systemic and we should be looking at systems and the big grand challenge at the end of the day involves system change, not changing energy decisions. So, in a way, evaluation is dominant kind of mainstream mode of operation is... This is why someone could say, "Are all evaluations SSH exercises?" Because I think most of SSH is kind of... Is excluded from the dominant mode of evaluation that we have in terms of projects, programmes, policies. So, there you go, that's my assessment of it.

Yeah, so some of the previous groups was basically arguing that all evaluations are of SSH and you're obviously saying that none of them are, or is that not true?

[overlapping conversation]

Well, it depends what you mean by SSH. Yeah, so but I think that when I talk about SSH, I talk about the criticals and interpretive social sciences into humanities, right?

Yeah, yeah.

So, I'm very careful about... Because I don't... Certain forms of economics maybe can get into this sort of area, but they have to... That's maybe ecological economics and things like that. But mainstream economics, classical etcetera would not entertain these sorts of things. And I think when we move into that space, we fundamentally change what we mean by evaluation. It becomes a completely different thing.

I quite like this idea. I'm [1:20:17] [inaudible] of paper. So, I'm thinking maybe we could... I leave you with your thought, before I start.

Oh okay. I was just gonna say, a different kind of this idea of what a different kind of evaluation would look like.

I could tell you in a minute, but you go first.

Well...

I've just recently written a paper on it so...

Have you? Oh okay.

So, I'll bore you with it a minute. So that means that you should go.

I was wondering what is the role of this SSH perspective in evaluation. I mean shall they be the representatives of the society? Are people's [1:21:00] [inaudible] over, only because I am a social scientist, I don't know how social issues are evaluated or if it's in good impact for my society, for my intercultural living together for elderlies; for I don't know what health systems, whatever.

What would the criteria be essentially?

Anyone could use social criteria in evaluation. You don't need to be a social scientist for doing the evaluation so it's... Social science method, for example, if you use a kind of grounded theory, I don't know a method for evaluation, would be this more iterative and circular thing but it's not about the content then.

Yeah... So, the process by you have participatory evaluations, you have evaluations that include users and other stake... Yeah, so you can kind of open it up, but I think the idea of social science and humanities kind of representing societies is again back into the service role. Often, we're drawn in, into interdisciplinary evaluations of projects, just to kind of engage, do the public engagement bit of the project.

What is the unique SSH...

[overlapping conversation]

Well, I don't think that changes the paradigm. What changes the paradigm?

An idea we had in other discussions about evaluation or... Not really evaluation, but in the development of any [1:22:37] [inaudible] policies maybe. Is to involve... Like in Germany we have the case that when a judge has two normal people as a side judge. And you get off your work and you get paid for those days and you are judging with a lawyer, a case, for over weeks. And it's kind of not really voluntary service for your society, and you could do the same in political decision-making processes for policies that at each table, for example, at the [1:23:19] [inaudible] steering group, they are all these governments, but each government person needs to have a working group in his country where citizens are part of. And not organized citizens, but just normal citizens. And they are chosen by... I don't know the English word... You could have a box and all the names, and which is drawn so...

Lottery?

Yeah, like a lottery. And you decide that you become part of the lottery, but you don't know when and how, but if you are chosen you have to do it. And then for five years, you are part of this working group, and you work policies together with the policymakers and politicians and things like that. And I think this would be a more systematic changing of how to develop policies and then also they come up with the evaluation criteria they think are important. I don't know what an elderly wants to... For him it's a good impact. I don't know it.

Yeah... I mean, this again, kind of speaks to the idea about opening evaluation up to. I think inclusive co­production, participatory is certainly something that I had in mind about the different kind of evaluation, but there's a couple of other things that... Going back to this point about beyond decisionism and thinking about systemic kind of challenges but also about inter­connectedness. Evaluations love to just have a very tightly bounded kind of space to the evaluation and you can be very kind of clear... Like I say, in terms of behaviour changes like, "oh they were in this programme or this information, and this change happened."

It's like setup, the whole kind of situation's set up to show the change, whereas what's actually happening in the real world is all of these things are kind of interconnected and interrelated with each other in wider systems. So, attending to kind of multiple ways in which things are productive and not having prescriptive criteria, or very narrow criteria by which you're judging things, being open to emergence and surprise. This then relates to the notion of experimentation or collective experimentation. So, this is rather than kind of doing a policy, it does something and then you evaluate it, the idea that we're all kind of in it together living a social experiment around energy transitions, whatever it is, leads to a very different way in which we're trying to judge or evaluate, which is a kind of continual, ongoing interrogation of the things that we're in.

And this is a very different kind of disposition when it comes to evaluating, and it means that we're not just obsessed with kind of policy­making then, we're interested in all the different... There's so many experiments going on. We're doing them in our households, we wouldn't call them experiment but we're playing around with stuff. I don't know, you probably got a new smart device, that's an experiment, bringing one of those into the home. So, you can have a very kind of distributed, informal kind of sense of evaluation in that sense.

The other thing that I would say is this linear cause effect kind of model, we only tend to evaluate good things that happen as a result of kind of it... So, if you look at the whole notion of impact of our science, do we ever... Well, actually, we do a little bit in the European context, in Horizon 2020, but hardly ever do we say... We always assume the impact's good, the impacts that we're having. If you flip that chart over, let's flip that chart over, on the left-hand side, there is something about... Or it's easier to gain impact, so you may have more impact internal in house. But you could argue that you're upholding a fundamentally unsustainable regime. Right?

Yeah.

We need to think about whether evaluate the downsides as well as... So, what might go wrong? And this is about anticipation as well. Not wait until after the event to evaluate. You evaluate almost before you start. And you say, "If we do this, whatever it is, energy intervention; What will it bring forward?" And then you adjust the policy or the innovation or the social engagement programme or new form of democracy, based on this anticipation about possible downsize feature... This is called responsible innovation in Horizon 2020 speaking, in European research communities. But I don't think this is part of the standard evaluation. Evaluations are normally done to rubberstamp, or to justify the legitimacy of something that's been done by an authority.

Yeah.

And this is what you were tapping into when you were talking about consultants, independents and all of that stuff.

Yeah.

The idea of critical evaluation or critical... Evaluation that challenges, that opens up the politics as well, is something... So then opening up is an important kind of, again, touchstone. Rather than closing the evaluation down, opening up and when your citizens come along and say, "Hey, this isn't about that, it's about this." "This isn't about energy. This is about basically the future world I wanna live in, and I think this, this and this is important." Rather than say "Sorry, that's out of scope, that can't fit in our evaluation," you say, "Okay, this policy is screwed up because it hasn't actually met the needs of the citizens that it's supposed to be serving." Do you know what I mean? That would involve a... A very open...

It would look very different... Yeah.

Yeah, so there you go. There's my point...

Manifesto. [chuckle]

If you allow for social science evaluation, how I understand it, you need to question basically the whole foundation of the European Union's language right now, like the buzzwords of circular economy and all those things, they don't work, and they have nothing to do with sustainability and the sustainable lifestyle and so on, and it's too political, for the next decades it won't happen. This is another thing, if you want to evaluate social change, it takes decades. And the assumption that you can evaluate something like social change in the short term means that you don't understand social change and its parameters. On the other hand, side, if you have a long-term goals, like you want to cut our energy consumption by blah blah blah, nothing happens now. And those are... Maybe also there's tensions between...

So long term, there's definitely... Yeah. Long­term definitely comes into it, which I think comes back to another point which I'll probably make later in the [1:30:36] [inaudible]. So, in this space, where I've done a lot of work is around how you evaluate basically societal engagement with say energy or climate change or whatever. And there are a million and one frameworks that basically before you do the engagement, it says this is the way it should be done, or these are the criteria of success, and then you do it and you judge it against those criteria, but they're normally very, very specific sort of criteria. So, a lot of what I just said is this kind of new... We've tried to develop a framework that doesn't pre­judge what the good qualities of the thing that you're trying to evaluate are, so it's all about kind of being reflective about how you're kind of framing the thing that you're trying to evaluate, inclusive, experimental.

These kind of things come into it, which are things that... Yeah, being concerned about what you've left out, for instance, rather than only just reporting the evaluations and... Do you know what I mean? It's like, what I find with all of this engagement stuff is, it's all about inclusion, but every engagement process that goes on is exclusive, you have to exclude some people and things to include... Do you know what I mean?

Yeah.

You can't every person or everything in the world, whatever those things are. But the thing is to...

But we don't ask evaluatively what is excluded, and so I think you can change the criteria or the questions that you're asking about a social science, humanities, can play a role in resetting the terms by which we evaluate things.

Also learning by mistakes.

Yeah, learning by doing, that's the experimentation.

And seeing mistakes as valuable for doing it better next time and don't expect that you can do everything perfectly in the first...

Yep.

And that you need those failures to become better.

Yep. So that's the experimentation and kind of learning by doing. And this would then... But to probably do that, you need this openness and transparency, and you need to get at the politics, 'cause often things that go wrong are not reported because then governments would go out, not be in power or companies would not make...

And you don't give money to them, again?

And this commercial confidentiality, so again, power and politics cuts through evaluation, and for evaluation to... A different kind of evaluation would also attend to power and politics in terms of what is shaping and framing? I mentioned it in the first group, but it's quite interesting that we see it and we say 3% to 4% of research funding is social science humanities in Europe. But it's like, how did that come about? That's about power and politics, right? It's about disciplinary power and politics, it's about... Do you know what I mean? And to say that this is all about kind of hard­nose facts and kind of, let's get to the kind of evidence now. Do you know what I mean?

Yeah, the thing is their... Along the administrative structure in the European Commission and executive agencies, they play a role in how the work programmes are made and how they take up evidence and things like that, and they are their own universes of knowledge, with their own agendas, and those administrative things, they are huge, and have a huge impact on what is written down at the end because all the steering groups, and all those working groups with members that in the end, there's a commission officer [1:34:19] [inaudible] sitting at the desk and writing down what is in the working papers. And there's no one in the end saying "Oh, I don't want that."

Yeah.

It's... Yeah, they are quite powerful, and this is also underestimated.

Comes a self-re­enforcing thing. Non­reporting of things that go wrong. So, in terms of the examples, I think that...

Has anyone done this?

Well, I think the STEPS Centre in Sussex, and [1:34:47] [inaudible] colleagues down there, [1:34:49] [inaudible] I think their work is aligned this way. We try and do it in the [removed for anonymity purposes]. I think Mathias Gross has done a lot of work, he's in Leipzig, on collective experimentation. There's a lot work in SDS on collective experimentation, there was a European report, "Taking the European Knowledge Society Seriously" in 2007. If that report was just copied out for this workshop, it would provide a lot of recommendation.

It's really great one, yeah.

Yeah. It's just the SDS guys saying... I think responsible research and innovation provides a lot of alternatives to the traditional evaluation frameworks. I think the demand centre, the one that was, not the one that is now, Elizabeth Shove and Gordon Walker's outfit. They did some quite nice stuff about policy into practice, which I think also can be informative. About taking a practice-oriented approach to starting to think about policy. So, they were a few of my examples.

And I think there's a long list of failures, which could be...

Yeah? What, as in bad examples?

Failures of...

Yeah, that's the thing, not bad examples, but like you've done evaluation, and you realize what didn't work, but you don't report on it. So, if you would make a look on failures then maybe you could improve your...

Yeah.

Yeah, I've been... I always say this at most events that I go to now, but the impact agenda in universities... It's across Europe as well, you're supposed to have impact through our research on the economy or policy making, that's what we're talking about, evidence into policy, and in the UK, we have these impact case study as part of our research excellence framework by which we get judged, in terms of our research quality. And all of those impact case studies only present the positives about the impact cases, so I'm just wondering whether we should just have some case studies of bad impact, like the internal combustion engine leading to emissions, which then led to climate... It was good, it gained mobility, but it also had a down... Do you know what I mean?

Yeah.

Asbestos, DDT, nuclear.

And if you look at all the...

They're good and bad, right?

The smart things, like smart is the new solution for everything. And if you think about how much energy we need for all these servers and all the thing...

Cloud.

The next generation will say "How stupid have you been?"

Yeah. And it's being presented as a positive environmental thing half the time.

But the impact case study of one of those things would say, "We're creating more efficient, connected... " I'm being quite jokey with this.

[background conversation]

[laughter]

But the point stands that we only tend to accentuate the positives, because then that gets us the funding and gets... Do you know what I mean?

Yeah.

And this plays into the evaluation process.

I had a talk with the project manager, and she said they try to include a critique on circular economy in the report and the policy officer said they're going to cut it out, because unless you don't get funding anymore. It's...

It's real.

Yeah.

It's not just like...

Yeah, it's closing.

We are not just making it up.

No. [chuckle]

No, and everything sounds alright so far, and...

There we go.

Thank you both very much, that's really interesting and...

It was interesting.

And yes. I better turn that off.

Let's come up with a different kind of evaluation...

I like it, I like our little manifesto.