Interview on Evaluation in Informal Science Education: Joshua Gutwill

Interviewee: Joshua Gutwill, Director of Visitor Research and Evaluation, Exploratorium
Interviewer: Alice Fu, SK Partners
Note-taker: Lisa Peterson, SK Partners
Date and Time of Interview: August 1, 2013, 10:30am to 11:30am (Pacific) (interview 2 of 2)
Location: Exploratorium (in person)

As part of our efforts to understand current evaluation issues in informal science education (ISE), we conducted interviews with leaders in the field. We purposely selected a sample of individuals who could provide insights from a range of perspectives; collectively, they have experience with ISE and ISE evaluation as practitioners, evaluators, researchers, funders, and institutional leaders. Several participants generously agreed to share the transcripts from their interviews.

Please note:

- These are transcripts of oral interviews, *not* polished or written remarks prepared for publication.
- These transcripts have been edited for clarity, brevity, and ease of reading. Participants were also provided with the opportunity to remove any potentially sensitive material.
- The views or opinions expressed are solely of the individual interviewee and do not necessarily represent those of their affiliated organizations.
- We intend for these transcripts to serve primarily educational purposes. We believe that others may benefit (as we did) from the rich insights provided in these interviews.

Interviews were semi-structured: we used a protocol that ensured asking key questions in a comparable fashion across interviews, but there was ample flexibility to allow for interesting and unpredicted turns in conversation. The coverage and order of questions varied across interviews. Interview topics included but were not limited to participants' views on evaluation uses, methodologies, "best practices," and challenges. Interviews were conducted in-person or by phone, and each lasted approximately 90 minutes.

In these transcripts, the following conventions are used:

- Initials indicate who is speaking. Blue text is used when interviewer is speaking.
- *Italics indicate paraphrasing or researchers' comments/interpretations.*
- 'single quotes' indicate hypothesized thoughts or questions; e.g., And I asked 'what have you had done before? And what did you think of it? And what do you need?'
- - single dash indicates an interrupted thought or change in thought; e.g., It's just been - I was just so happy to have had that opportunity to work with them.
- … ellipses indicate overlapping speech.
- [brackets] indicate non-verbal observations and other clarifications added by SK Partners.

A Note about this Interview:

This is the second of two interviews that SK Partners conducted with Joshua Gutwill. Parts of both interviews focused on the Exploratorium's Group Inquiry by Visitors at Exhibits (GIVE) project. For more information, see the GIVE project website:

http://www.exploratorium.edu/vre/visitor_research/give//index.html

AF: We spent a lot of our time last week talking about design and choosing the right questions and the right methods; we talked about APE and GIVE, as well as other projects; so we'll go back to some of those projects. I think we mentioned last week that we want to focus this conversation around uses and stakeholders, if that's okay.

JG: Sure.

AF: And, towards the end, depending on time, I have a few more general questions about evaluation.

JG: Sounds good.

AF: Do you have any questions before we dive in? *[JG has no questions].* Alright, let's dive in and go back to GIVE. Can just describe for me who you saw as the primary stakeholders or users for that project?

JG: There are a couple different stakeholders. It was a research project, so, I think one set of stakeholders were, and are, other educational researchers interested particularly in informal environments. The big question there was, can we, and how can we deepen inquiry in informal environments--scientific inquiry--and sort of sharpen it, get people to be a little more intentional, a little more systematic. There, what we were looking at particularly, the difference between Juicy Question and Hands Off was a pedagogical difference, about spontaneity in the Hands Off vs. more structure in Juicy Question. There was a theoretical question there around the kind of pedagogical approach that would work well in an informal environment and whether it was different for family groups and field trip groups. It was a learning question and we thought that educational researchers and educators would be interested in that.

And then there was more of a practitioner stakeholder group. That is people who actually develop and implement programs for groups in informal environments. People who might want to do this with school groups or want to do this with family groups.

Those two primary stakeholders were well represented at the beginning of the project by the leadership team. Sue Allen was the PI and she is a researcher, and then you had Erin Wilson, who at that time was our Manager of Public Programs as a co-PI. So the idea was that she would bring expertise in kind of, what is feasible, what would actually work, to kind of keep the research honest. You don't want to develop and test something that is completely ridiculous to apply in a real environment. But then she would also help implement whatever we found afterwards. She would put that into public programs. She left the museum halfway, partway, at some point in the project, I forget exactly when, as did Sue. But I took over the PI role and

represented the research piece of things. We didn't have someone to take over that practitioner role. As we get to dissemination, which I know you are interested in, I think that was an oversight and I think maybe has led to less uptake within the institution than I had hoped. Although, as we'll get to it, there has been some uptake, so that's gratifying.

And then of course the funder is a stakeholder. They want to make sure we do good work, and want to be sure we disseminate our work.

And then, you know, just the power structure here at the Exploratorium is a stakeholder. They want to be sure we're doing good work and that we are moving the field along, you know, helping to move the field. So there are stakeholders there as well.

But I think the two main ones are the educator / educational researchers and practitioners.

AF: And they are, the educators, because it is a research question about learning and the practitioners because they need to develop and implement programs for different groups.

JG: Yeah, so we ended up doing kind of two things. One was offering people an actual, simple thing you could do in museums called Juicy Question that has been rigorously tested and found to have these effects, at least in a lab environment. And we've also learned some things about how one might develop a program, a simple program, and we share those. So for example, we published, 4 different pieces around this project. We published 3 journal articles and a book. One of the journal articles is in Curator, and that is really targeted at practitioners. It's really about- it was even before we had the results- it was, how did we develop these programs? It talked a lot about the formative testing and the formative development that we did. That was very much for them. The book was also for practitioners, although it has all the research in it as well. But we wrote it in everyday language, or at least we hoped to not use a lot of jargon. And then we wrote two journal articles, one for Science Education and one for the Journal of Learning Sciences, those were really pitched at that educational research / education theory folks.

AF: Ok, cool. So you anticipated the next question, which was around these different dissemination and communication strategies. Are there others you want to talk about?

JG: We wrote a bunch of things and we've given a lot of talks. If you want I can actually try to count up all the talks that I've given and I can probably find out what Sue has given. It's probably on the order of a dozen to 20. If you look at internal, external, conferences, informal, invited, this and that. It's probably something like- between the two of us it might be 30. We've given a lot of talks. I can actually find out for you. A fair number. We promoted it a lot.

AF: And you said some of them were internal within the Exploratorium?

JG: Yeah. We did brown bags within the museum, I've done conference talks. UW invited me up there to give a talk about GIVE. The science center NEMO in Amsterdam invited me there to give a talk, that was super fun [*JG, AF, LP laugh and joke about how fun it would be to go to Amsterdam*] all expense paid trip to Amsterdam? Like yes! So things like that. I've been a guest lecturer at the JFK program in Berkeley for a bunch of years and talked about GIVE at every one of those, once a year. Sometimes there is a SF state class that occasionally the professor contacts us and, says 'can I bring my students in and you guys give them a little song and dance?' We've done that, talked about GIVE, it's been promoted in a whole bunch of different ways.

AF: Cool, okay! And then there is also the website.

JG: Yeah, and then we have the website. [GIVE] has its own website, plus the papers or preprints of the papers, and other reports. There is a VRE, a Visitor Research and Evaluation website that has all of our reports on it. So the reports are there, plus there is a link to the GIVE website; both of those. Um, *[JG pauses]* I'm trying to remember, I think, yeah.

AF: So that's a lot of different kind of arms. Do you know that going in? That you are going to tackle it in all those different ways?

JG: We did promise in our proposal that we would write papers and a book. Which was really good cuz it meant we had a budget for the book.

AF: Yes, okay.

JG: So we did know that. And I think we said conference presentations. But I don't think we went in thinking we will go to museums and give talks or we will go to universities and give talks. Oh, and I think we gave a talk about it over at Berkeley too. Anyway, I don't think we promised that in the proposal. But as our work has gotten a little more known, and also when you publish, that's when people, like UW read about it and go like, 'hey, why don't you come here to talk to us about it.'

AF: Or Amsterdam [*AF and JG laugh.*]

JG: Or in Amsterdam. Yes, which has led to- anyway, we'll get to uses.

AF: Yes, we are kind of – we are moving from who are we targeting, how did we communicate to them and then what did they do with it; how were the findings used?

JG: Yeah, so one of the talks I gave on it was an invited thing down to Cal State Long Beach. Jim Kisiel is down there. He had a class going and he invited me down to talk about GIVE, and so I did. And then he sent me an email, maybe a month or two ago, at the end of the semester. I can ask him, if you want to see it, I can ask him if it's okay to forward it to you, I'm sure it will be. Basically he tried doing Juicy Question, with a little bit of adaptation, tweaks, with his students, to give them a sense of 'what is science inquiry like in informal learning?' So he used Juicy Question in his class, and said it was fantastic. And kind of had all these things about what worked and why. He was very excited.

AF: Cool. I would love to see it.

JG: Okay great! And then after the Amsterdam talk, a TV producer came up to me, [laughs and jokes] and said- I was like keep your head small, keep your head small- and said, 'Juicy Question sounds amazing and we are creating this TV show, where in each episode we do some cool experiment.' It's not exactly like Myth Busters, but it's got a little of that. The idea is that you the viewer would go and try some stuff and do an experiment. There is even an iPad app / site that goes along with it that you can get and then it supports you in doing things. 'We would like to use Juicy Question as our structuring approach for the experiment part, where they actually do things, like get people to ask Juicy Questions and do the experiments, and share discoveries.' And so I said, 'that sounds like a good idea. I would be in favor of that.' *[All laugh]* So it's been minimal input on my part. She sent me a storyboard for the app, 'here's what this might look like.' And I gave some input and feedback on it. I don't know where it is now, they are working to implement it. What is so sad, even if I could watch it, it's in Dutch. [laughter]

AF: You don't understand it. You need a translator.

JG: Right. But it's very exciting, it's such a cool thing to hear that.

AF: That's really cool. Who would have thought?

JG: I know, really, totally crazy.

LP: Unanticipated consequences *[all laugh]*

JG: yeah

JG: I don't know if you would call this implementation of GIVE, but afterwards, in Amsterdam, one of the people there is named Tessa van Schijndel, and she was a student of Laura Schulz' at MIT, she was a post-doc with Laura Schulz at MIT, maybe two or three years ago. She actually came out to the Exploratorium at that time and met with me, with her colleague from the

University of Amsterdam, Maartje Raijmakers, and they ended up being the people to invite me out to NEMO. But they met with me here, we talked about APE, we talked about GIVE, and they gave a brown bag about their work to the staff, which was cool. So, I don't know if this is like GIVE, but Tessa and I are now exploring potential joint research proposals together. She is a developmental psychologist, so she is really interested in the younger children, but we're both interested in inquiry and how to study and foster more systematic inquiry. Again, I don't know if that's exactly like GIVE adaptation, but…

[JG reading from his computer screen]

Also, I haven't checked up on this, but I did talk to Sue, who is out here right now as a matter of fact. Sue Allen.

AF: oh is she? We interviewed her over the phone.

JG: Yeah, she's here. I think she is here for another week or so, if you need a follow up or if you just want to put a face to a name or meet her of you've met her.

LP: All of the above.

AF: I met her once long ago. Lisa hasn't though. That's cool.

JG: Anyway, she told me that the New York Hall of Science was 'very excited' to use Juicy Question on their floor, and I don't know in what capacity, if that's with school groups or-. And she said ask Peggy Monahan.

AF: We have an interview with her, Peggy, scheduled for end of August about Secrets of Circles exhibition down at CDM.

JG: Oh right. Excellent. Maybe you could ask her about this. Or do you want me to follow up?

AF: No I'd be happy to ask her

JG: Again, I don't know, but that's what she said. The other thing was the people at the EcoTarium, Alexander Goldowsky, who I think has just recently left [EcoTarium]. EcoTarium is in Worcester, Mass. I think he has just left to go to the MIT museum. But he and Betsy Loring at the EcoTarium came out here a year or so ago and met with me and were kind of picking my brain about GIVE and Juicy Question. We talked about doing a proposal together that would be more of like an implementation proposal for small museums to start using Juicy Question. And we've all just been too busy, honestly, to deal with it.

They invited me out for a talk at the Acton Discovery Museum a couple years ago, and I talked about GIVE then. And subsequently, they have invited me to come out and give a workshop, which I'm doing at the end of September, about open-ended exhibit design and I'll probably talk a little bit about, well, I don't know how much we'll cover GIVE, it's more about APE stuff. It's kind of all related work.

AF: The workshop will be at the EcoTarium?

JG: It will actually be at the Acton Discovery Museum.

AF: Oh okay.

JG: So Erik Toberson, he's a developer here, and I are doing that together at the end of September.

AF: wow

JG: So lots of…

AF: Lots of stuff, yeah.

JG: And then, in terms of internally, I was saying before, I think it was a tactical error to let Erin go and not try to have someone else step on to the project. The truth is we were so far along in the project. I think she left either while we were collecting data or after we had already collected data and had analyzed it. I mean, it was late, it was way past the point of creating the programs, you know, the Juicy Question and Hands Off. So to have someone join, what would they do? So I met with the Explainer folks. And I think, are you talking to some of them?

AF: I was planning on reaching out to Anne Richardson

JG: Yeah. So Anne would be good and Liana Crouch. I met with them and showed them the GIVE work and Juicy Question and the results. And Anne was, they were both very interested. Anne was finishing her dissertation at the time, so was really interested. Well, busy, yeah. But she was almost the perfect audience because she was a practitioner turned researcher; she wanted it all. She was very psyched and thought it was good work.

When they actually started to work with the Explainers, and think about 'what would we do', they felt like 'this echoes what we already do.' Of course because we went to them to say 'what do you do?' and then we tried to distill it down. Their programs are really about professional

8

development for high school kids or youth development for high school kids and sort of college age kids. It's really about empowering them and having them grow and learn. They don't want to give them, you know, 'here's your tool, go use this tool.' What they like to do is have the kids in the program co-create things. Let's do a new kind of demo and then they develop it together. What they felt was, why don't we show them the Juicy Question game and have them practice it and play it, but say to them, 'this is a tool in your toolbox. You can use this however you want.' So it was very open with them. You should check with her about where has that gone, what are they doing with that now. I'm not really sure. But having since moved to this new site, we are now offering paid tours.

AF: Right. You mentioned this really briefly last time, over the phone.

JG: yeah, so there is a group led by Ken Finn and Kurt Feichtmeir here who are putting together these tours, figuring out what they ought to be. I mean they are already doing it. So, I don't know, 6 months ago or something, before we moved in here, I met with Kurt and said, 'I know you guys are doing these tours, I just want to offer, if you want to do something around Juicy Question,' it could be a 'learn how to use the Exploratorium more deeply' or something, because that's kind of what it does, is it gets you more deeply into the exhibits and it's designed to be used at any exhibit and so on. I can imagine a tour where you teach it to them and then you scaffold them at a couple more and then you turn them loose. It could even be a short thing, a half an hour, and you learn how to use the Exploratorium. He was interested and they have been looking at that stuff. Well, apparently, they have started up a reading group with their tour guides and they are not offering a Juicy Question tour yet, but they are doing a reading group and next week they want me to come and talk because they are going to read one of the papers about GIVE. And so next week he wants me to come and meet with them and answer questions and talk more about what GIVE is and Juicy Question and all that. So I think it is getting taken up in some form or another, I don't know what form it is yet, by that tour group. So there is something there.

AF: That's cool.

JG: And then, there are also citations. I actually did go and look cuz I was curious.

AF: Oh you did?

JG: Yeah.

AF: What did you find?

JG: For the three articles and the paper, there have been so far 41 citations since 2010.

AF: Wow.

JG: So 41 articles have cited the work.

AF: Yeah. That's pretty awesome.

JG: Is it? I don't know. *[laughs]*

AF: I'd be excited about that.

JG: Yeah. Well, considering two cited my dissertation, that's pretty good. *[laughter]* Yeah, so it is getting used.

AF: Cool, ok, that was great. Before I move, I'm going to ask you to now do this chain for a summative evaluation project. So before I do that, do you want to say anything more about uses and dissemination on GIVE?

JG: *[pauses]* No, I think that's all I can think of.

AF: Okay. So I do want to compare this [GIVE] to what would be more of a typical summative evaluation experience in terms of those three things. Who are the stakeholders on a summative evaluation project, how do you communicate to them, and then how is it used?

JG: And you know that I haven't done summative evaluation project, Well actually I have, but as a consultant outside of here.

AF: Yeah, so think about summative evaluations that are done for you.

JG: For us. Ok, got it. Yup.

AF: And let's walk through that. I don't know what is easiest for you, you can talk in general or if you want to pick one and do that more specifically, I'll let you decide.

JG: why don't we start general and then if something comes up--

AF: And if you have examples.

JG: So who are the stakeholders?

AF: Yeah. so, you listed 4 main groups, well two main groups and then two other groups. For GIVE, you'd said education researchers, practitioners, the funder, and then kind of internally the power structure of the Exploratorium. How would those compare to stakeholders on a summative?

JG: I think on a summative, it is probably the same folks, but the priorities shift around a lot.

With GIVE it really felt like Educational Researchers, Practitioners maybe a little bit behind them, and then Funders a little bit behind, actually a fair amount behind them- because if we publish in a peer reviewed journal they are going to be happy- and then the powers at the Exploratorium.

Whereas with summative, I'm not sure who has the priority, but it's funders- a lot of the feeling is 'we're doing this because it is mandated by the funder.' The second one is, not the powers that be at the Exploratorium, but the team. And those are probably same priority. The team wants to know, like, 'ok, how did we do?' Not just the team, but the team and the people the team reports to. In our case, it would be like Tom Rockwell. We want to know we did a good job, and we want him to know we did a good job. And we want to know where things didn't work.

One of the nice things about the Exploratorium is that there really is a culture of learning here. There is total support for the idea that you can learn through failure. In fact, sometimes you learn best through failure. So failure is not the worst thing in the world. People don't lose their jobs because something didn't work, here. A real problem would be you failed and you don't try to figure out anything about why. That would be a real problem. But not hitting the mark is ok as long as you try to understand what is going on there. So anyway, those internal folks are really key stakeholders. Funders and internal folks are the most important stakeholders.

Other practitioners and educational researchers, they are so low on the priority list. I think occasionally something happens and it's almost serendipity, where a summative evaluation learns something that can actually speak to theory or *[laughs],* you know, something a little beyond practice, theory light. But mostly not. Because that's the problem, they are not very generalizable. That's exactly the constraint.

AF: Right.

JG: And maybe there is a way to make them more so. I don't know, it's hard. And then practitioners outside of here, if you can structure it in a way that tries to get at something that might, design-wise that might--not theory-wise, theory is really tough--but design-wise that gets at something more general than just 'is this a good project? Did we make a nice exhibition?' I think you are on to something that other practitioners can utilize.

JG: Well, so going back to that APE example of Selinda Research, having this finding of 'try to keep your visiting group together because people will go deeper if they are together.' I think that is something that practitioners can use throughout the informal world. And I think Deb Perry's approach- she has a naturalistic inquiry approach and she also has a theoretical framework- I think having a clear lens and utilizing that lens in a clear way can mean that other practitioners can say, 'well I'm interested in that lens too, I'm interested in that, in thinking about the intersection of design and motivation and learning outcome,' which is kind of her thing. So let me look at Deb Perry's summatives and see what she has found using that lens. So I think there are ways to do it. But again it's a lot of work, for the practitioner, it is a lot of legwork. We post them all on informalscience.org, and now it's CAISE's informalscience.org, which is definitely a good step, so they are at least available. It used to be, you would call up Deb Perry and say, 'I heard you did the summative for x. Can I have it?' And she would say 'No, it's the property of x. You can't have it. Call x.' Or, if she was nice, she would say, 'I'll check with x and if they'll let me, I'll give it to you.'

AF: Right.

JG: Now NSF mandates, 'no, these things have to go up.' So at least they are available, but it is a lot of work for somebody to say, 'I'm interested in how visitors think spatially.' And then, 'well, gosh, have there been any projects that have looked at spatial reasoning?' And if they are lucky, they'll stumble across the Geometry Playground summative evaluation. Or if they are lucky, it has been tagged with 'spatial reasoning' and they can find it. But there is nothing in the title of Geometry Playground that it's about spatial reasoning and that part of what they looked at in the summative was to what degree and to what extent are people reasoning spatially. That's not in the title. So how would you know? So if you are looking for something general like that, it is tough. It's more like, 'we're doing something on geometry, has anything on geometry been done before?'

AF: Right and search geometry.

JG: And then you look and you go like 'oh, but we're doing proofs and they did spatial reasoning, that's not useful.' So it's hard, it's hard to really make use of the stuff. Even if we did proofs, then you would have to look through the whole thing to see if there was anything here that's applicable to my work.

I think in regular journal research, part of what works really well is we cite each other, we build on each other's work, so you can tell within a page or two, 'what's my orientation' even if I

don't have a clear theoretical framework, at least, what is my orientation? Am I in cognitive science? Am I socio-cultural? Am I, whatever. You can start to say, 'ok, I can see how this led from that. Ok, Josh and Sue did this work. So what did it build on, what's its antecedent?' That's not so much true in a summative evaluation.

Deborah Perry is a nice example, where she has a lens, she has a framework, but a lot of other summative evaluators don't. They kind of come in, they are very open, they are like 'what were your goals? What were you trying to do? I have kind of a toolbox. I'll try to get at whether or not you reached your goals.' But it's not like from one summative to the next you can see a through line very clearly. To be fair I haven't actually done this, where I've said, 'let me choose a summative evaluator and see if there is a through line.' Maybe for lots of summative evaluators there are through lines, but I'd be surprised to find that.

AF: Yeah.

JG: It makes it hard for other people outside the funder who gets to say 'check, we gave them money and they used it appropriately.' And the internal staff to say 'great, we did this' *[gestures making a check mark]* or 'oh, that didn't work so well, let's try to understand that better.' And that's something we try to do, even in our summatives is, stay close to the evaluator so that as they are getting data and they are starting to say, 'this seems like it is problematic.' We can say 'oh, can you follow up on that with-?'

Here is an example of that with Geometry Playground. We had it set up here without really any walls. It was classic Exploratorium, like a bunch of exhibits. It had kind of one wall in the front with an entryway. Which is a little bit funny, right? Like an entry with nothing behind it. Anyway, there was an entryway that said 'geometry playground' and it said stuff about 'learn about geometry in a fun way,' and thinking spatially or something, I forget. And one of the early results was, at Selinda, they were saying, people don't know the name 'geometry playground' because that's the only place that they see it. And we are seeing lots of people coming in to the exhibition through an open side or an open back. We're seeing people leave that way and then come back. There isn't like 'a flow' through. We liked that because that's how we are at the Exploratorium, but we wanted people to understand that it was called 'geometry playground' because we felt like that was good framing. In just two words you get a pretty good frame for what it is. So we asked them [Selinda]- they were already planning to study it at other institutions 'cuz it's a traveling exhibition, so they were going to look at it at several other institutions. And we said explicitly, 'Can you look at this issue in other places where they are going to have it in a room, or they are going to have it walled off in some way, people are going to have to walk through that archway to get in. Can you look at that?' And they did. So that is an example of where the team stakeholders get to ask.

AF: So I had a question about that actually. In this example you had an evaluator who was open to your input and you were able to put your question and your issues of concern into their study. You talked about other examples last week, about others who were less receptive to your input or your having this participatory role. Is there a relationship between your level of participation and level of usefulness of a given evaluation, or is it not that simple?

JG: Hm. *[pause]* I don't know that I have enough experience with a broad enough array of evaluators to really say for sure; but my hunch, my feeling is that, yeah, the more- and I think this is true of any evaluation- the more the stakeholders are involved in helping the evaluator understand what the questions are, what the issues are, what the problems are, the better, the more useful the evaluation will be to those stakeholders. So I think, yeah, if we can participate more-

That example I gave last time, I wasn't the person interacting with the [evaluator], it was actually Sue interacting with the evaluator. It was less about what are your questions and more about what are your methods. So Sue and the evaluator got into a bit of a tussle, a short-lived tussle about 'are you going to do it this way or are you going to do it that way.' Finally the evaluator said, 'I want you to stop. This is my show and I'm going to choose the method.' And we all kind of talked about it and agreed. It actually is their purview to determine what the methods are. It is outside of our purview.

And actually, this is a little bit of a tangent, but summative evaluators are not covered by our IRB process, they have to get their own, and specifically because we cannot take responsibility and don't want to take responsibility for their methods. So their methodological choices are really their own. They have independence in that way. That's the whole point, I mean, that's why NSF says it has to be an external person. They should have independence; they should determine what the study is, how the study works.

AF: There were so many things in what we just talked about that I want to follow up on. I'm trying to, like- Oh, well, since we are talking about relationship with evaluators, how do you generally choose an evaluator? What do you look for?

JG: Well *[pause]* That's a good question. Most of the evaluators, all the evaluators we use, we know through VSA meetings or just through the field. You start to get a sense of their work and their approach.

One guiding principle, that was really Sue's principle, and I've tried to maintain it, is to try to have diversity. Don't just always go back to the same people you always use. There are even people here at the Exploratorium, in other departments, who use the same person, over and over and over again, or the same group over and over and over again. And that's okay, but part why

we want- So I guess this is getting at your question, part of why we want to use different folks is we actually want to learn from them about their methods. So it's fun to go to Selinda Research and learn all about naturalistic inquiry. And then go to George Hein, where he's got a very constructivist, almost a radical constructivist stance and he uses that lens to think about what people are doing, which is great. And then go to Minda Borun who thinks a lot about design, and go to Beverly Serrell who thinks a lot about how people move through a space. Just going to different people and getting their approach and their methods. On the projects I've been involved with, where I am the liaison to the summative evaluator, I always ask them, 'can you come give a talk to the team about your approach. Come tell us what you do and how you do it.' To just educate us and all the researchers and evaluators who are on the team.

AF: Interesting. Okay.

JG: We usually invite other visitor research staff in if they can make it. 'They are coming and they are talking, come listen if you want to.' I didn't know anything about naturalistic inquiry, for example, until Deborah Perry worked on the APE project. I learned a lot about it. So it's looking for a spread, it's looking for someone who is professional, who is sharp, who is honestly, easy to work with. Most of them are, because you can't really succeed in that work if you're not interpersonally savvy. It's a totally client-based profession. So yeah, those are some of the general attributes. But mostly variety.

AF: Yeah. That's cool. That's interesting. One of the things you mentioned was around the culture of learning at the Exploratorium, I'm wondering if you can talk a little more about that. In particular, I'm wondering about how evaluation and research findings play into decision making. Either, in general or with exhibits in particular.

JG: In terms of the culture of learning in general, it has been there since the founding. Frank Oppenheimer really epitomized that view in his work. He really was a tinkerer. He really supported other people in messing about and trying to either make sense of things by making things or making things to help other people make sense of things. He published a little book years ago called 'working prototypes' and what he says in there is 'our exhibits are never finished. They are all working prototypes. You can always learn something and improve something, add something, take something away. You can always learn more about how people interact with these things.' So I think that has been infused through the whole institution. As people come here they get enculturated. When I first started, there were literally people who knocked on my door and said 'hey, welcome to the Exploratorium, have you read this thing that Frank wrote? And did you know about this thing that we did?' There is no library of stuff you are supposed to go read, but there are people here, and it's from all different walks of life in the museum, who- myself included actually now- who think it is important to continue the traditions, and that there is a real culture here and it is important to keep that culture. It evolves and it shifts,

and it needs to, but there are some core ideas that are pretty strong and I think that is one of them; that really has been maintained. Anyway, so in general, I think, that's very much here. If you go to an all staff meeting and hear Dennis talk to the staff, you'll hear it in just the way he talks about things and says things, little things he'll drop, that everyone in the audience is like 'oh sure' but you would go like 'oh, oh, that's that culture of learning thing.' It just goes right by us because it is part of the culture. But you would notice it.

So, coming back to, 'are decisions made based on research?' I think it really depends on research and evaluation. First of all, in terms of formative evaluation, absolutely, decisions are made everyday at an exhibit level, based on formative evaluation. Sure. In fact a lot of our developers ask us for evaluation of x, y, or z, so that they can decide, 'do I do it this way or do I do it that way?' We also do a fair amount of audio and videotaping, even in our evaluations and sometimes developers just want to watch the tape. 'Let me just see what visitors are doing where I can really hear them.' One of the advantages of audiotaping is that you can really hear them, whereas if you are just standing on the floor it is harder to hear. Although in this new building it is definitely quieter, I've noticed a substantial change.

AF: Yeah, the other one was so cavernous.

JG: Right. When you're on the floor you hear lots of sounds, but there was like a deafening roar in the other museum when we were crowded. It was really tough. You had to yell, you had to raise your voice. And if you turned your back on someone, *[JG turns his back]* you could not hear, they would not be able to hear what you said at all. I haven't tried that, I need to try that on a busy day here. I don't think that is true here. Anyway, so developers will ask for it.

At a higher level, I think, at this building we've started to get into the game a little bit, we--VRE, have gotten into the game of doing evaluation that helps answer, like higher level questions. So for example, right now we are doing a study where we are out on the floor, I think it is like 4 days a week and we have interviewed now over 5,000 visitors since April.

AF: Wow.

JG: So we are collecting a boatload of data and it's all on iPad, so it just instantly goes into a database. One of the key questions is, 'at what point of crowdedness does the visitor experience suffer?' So we know how many people are in the building at any given moment because not only do we sell the tickets, but we actually have cameras. We have a camera system at the exits and entrances, so any time you move from a ticketed space to an unticketed space, there's a camera overhead that just counts you. It doesn't capture your image, it just counts you. So we know the flow in and out. We know at any given moment how many people are in the museum. And then we are looking to correlate that with a Likert scale question about overall experience from 1 to

10 at the moment that we conduct the interview. We also ask people about crowdedness and things like that and was there anything frustrating or confusing and sometimes people talk about crowds. So we are gathering this data to help the admissions folks decide at what point do we stop allowing people to come in to the building? Like is it at 4,000 or 5,000, 6,000? When does the visitor experience deteriorate enough that you actually want to stop selling tickets. So that is a study that we are doing. They are looking at our data now. We don't actually have the correlation yet because their camera system was malfunctioning until about 2 weeks ago. So we know how many tickets we sold but we don't know exactly how many people are in the building. So we're trying to figure out 'what is the right correlation to run?' But that is beside the point. The point is they are using that. We really haven't been doing stuff like, at that level, of whole museum work…

AF: Yeah, yeah

JG: …until really recently where there would be big decisions made.

At the exhibition level, however, there have been things, like we will do studies on how people move through an exhibition and that will lead an exhibition team to change the way the exhibition is structured or laid out. What else have we done? We have looked at immersive exhibits vs. table top exhibits, and the spatial reasoning, and so on, and there the results were mixed—some things are better at the table tops and some things are better at the immersives so it wasn't such an obvious like, 'throw away all the table tops, we should only do immersives from now on.' [AF laughs] or vice versa. So there was not an obvious thing to do. It was more like, if you want to do this, do it immersively, if you want to do that, do it table top. So that's something that people like Tom Rockwell take in and use for the next project.

AF: OK.

JG: But actually, that example I gave you, of the 5000 data points, I feel like that's a great example of a culture of learning,

AF: Yes.

JG: The marketing folks and the marketing admissions, business development people, they come to me and they say, 'I have a question, can you help me answer it?' And I say 'yes'. And we work together to do it. That's probably not super unusual. I'm sure a lot of marketing groups do that at other institutions. It is just an example of 'we want to learn, we want to know what's working and what's not.'

AF: I'm looking at time, we just have a few more minutes left. I want to be sure I ask this question. Moving to a broad question around needs for the field. What do you see as the most critical evaluation-related needs and knowledge gaps for the field of informal science education? In other words, where does this field need to go?

JG: Oh god. [AF and LP laugh.] It's funny because I was at a CAISE convening in, I think it was March or February or something, where basically this is what we looked at. Can I just hand you that, or at least just look back at the document before I answer?

AF: That would be fine, or if you are allowed to share it, you can always just send it.

JG: I don't know if I am allowed to share it. I could ask Jamie Bell. Actually, just give me a second, that might actually be worth just looking at *[looks at his computer, pulls up the CAISE document]*. Because I feel like we did come up with some stuff that's- Yeah, what we came out with was a broad array of different things the field could look at.

So, learning ecologies and change over time. Trying to understand how people move across these different institutions and experiences and informal spaces and how learning gets stitched together or not across these things; seems like it would be a really interesting and important thing.

Identity formation. How do people become really deeply interested, to where it is not situational interest anymore but it's like intrinsic interest in STEM. How do we foster that? How do we have people who say, 'I want to be a scientist when I grow up.'

Interest, motivation and curiosity. I think there has been a lot of work done in interest development, but interest development in the formal world. Which is kind of ironic given that interest and sparking interest is something that informals are so good at. There are a lot of people who say, in evaluations, you often get Likert scales of how interested were you in this or how many people expressed an interest. But I haven't seen a lot of actually developing interest and even taking, Hidi and Renninger's model of interest development and saying 'let's apply that in an informal context. What's happening for people? Are they actually shifting in the kinds of interest they feel?' So that seems cool. Motivation: Falk has this visitor identity theory, which I- and I think others in the field- I think that it is actually related to identity, he calls it identity with a small 'i' because he's trying to say, 'I'm not talking about who I am.' But it seems like it's very much related to motivation. Why did you visit today and how does that affect what your experience is in this space?

And then, understand learners and learning. Ok, people are doing that.

21st century citizens. This isn't something that I've been thinking a ton about, but really it's about digital age and technology and what they talk about 'mainline' as opposed to 'pipeline.' The people who aren't going to become scientists, but still have to know a lot about [science and technology], or at least if they were more informed about science and technology might be able to make different kinds of choices in terms of their purchases, in terms of their energy consumption, in terms of their citizenry, and voting, and activism and things like that.

Design principles to support ISE learners and learning. That is very much near and dear to my heart. This was about practitioners and researchers working together. And I think, looking for design principles is a great nexus point for researchers and practitioners to work together and that's why I think it also works well for evaluation.

Methods, instruments, and assessments. I think we could push a lot harder in our field for new methods. I mean we are; you go to VSA and there are people talking about new methods, but I think it would be great to do more work in that realm. And also, honestly, working more with statisticians and psychometricians and trying to beef up the rigor in evaluations- of the modeling and the statistical methods. I could use more support in that area myself.

There is another one here, capacity building, but I don't think that is really about evaluation per se. So I'm glad I had that list, because I would have thought of about 3 of those, 3 out of the 8.

AF: [laughs] You covered a lot.

JG: That is me and others thinking.

AF: I want to make sure we get out of here so you can get on with your day. Is there anything else you would like to share about your thoughts on evaluation before we leave you? You can always e-mail if you think of something else.

JG: I don't know what kind of picture I painted of the interplay here between researchers, evaluators, developers and practitioners. I just want to say, it ranges; there are some people, practitioners who absolutely want research and evaluation, they feel like it is a support to them. And that's what we are trying to convey. We are here to help you realize your vision and hopefully learn something together about how people learn. And then there are people who are either confused or hesitant. I think over the time I have been here, I have seen a big shift though from, when I first came, there was really a lot of resistance- I guess there was resistance, but there was a lot of skepticism, that's really the word. What is this and are you going to give me a grade? You know, is that what you are about? You are going to tell me if I succeeded or not. You are going to tell me if I got an A+ or a B-. Maybe we already talked about this. I can't remember

JG: I feel like there has been a big shift away from that. That is something. In general, I think good relationships in evaluation are based on mutual trust. And one way to really build that trust with practitioners, is to really underscore the role of the evaluation, is to try to learn something that will help you—you the practitioner. That's what I'm doing here. Even summative evaluations which are really about, not necessarily a grade, but really are about, 'to what extent did this succeed?' They can be framed that way. And they can really look at what's working and why, and what's not working and why, and not just like, 'well this is bad.' Evaluators range too, in how critical they are, and how much they like to throw criticism around, and things like that. The more we can frame things as 'I'm trying to learn something here, too, that hopefully will improve your practice as well as the learning experience for people.' I think that is more on the right track.

AF: Ok. Cool. Thank you. I'm glad you added that little bit. One of the questions we wanted to talk about, how to build relationships and cultivating buy-in. I'm glad you talked about that.

JG: Oh okay, sure. I couldn't remember if we had talked about it. I've thought a lot about that. It is so much part of the work here, is building that trust, and building those relationships.

AF: And how do you do- does it just take time?

JG: Yeah, so time is key. Understanding their perspective, really listening to them is important. As I said, having that orientation of 'I'm here to help you.' Especially, it's so easy to do when you are doing formative evaluation, because that really is what you are doing. And then, you know, sometimes what has really worked is bringing them into the evaluation process. Giving them a clipboard and saying 'here are some behaviors that we agree on to look for. Why don't you go out there and actually watch people, time how long does that person spend, here's a stopwatch.' Get them to do a little evaluation. That often gets them excited about it. And this might be a little self-serving, but it helps them see how hard it is. They realize, 'Oh, being systematic is not an easy thing to do.' There are a lot of issues that come up and you have to decide, 'Am I going to do it that way or that way? How do I decide if someone is at my exhibit now? When do I press start on my stopwatch? When do I press stop?' They start to realize, it is all gray and the devil is in the details, just like making an exhibit.

AF: Right.

JG: It starts to build mutual respect. A lot of times, when they don't understand the process, they just think, 'Well I could go out there and talk to people, what do I need you for?' When you say,

'Ok, well, let's do it together in a systematic way.' They start going like, 'oh, I see.' So there's a bunch of different things you can do. Super important.

AF: Sorry, I snuck that in there.

JG: No. It's actually near and dear to my heart, which is probably why I thought of it at the end. So I'm glad we're talking about it a little bit, because it's really important.

AF: Thanks, Josh!

JG: Sure, thank you!

AF: This was awesome. I am really glad we got the time to sit and down and talk through these issues. We are so grateful.

[END INTERVIEW]