JB: Welcome to Listen to THIS, where we have conversations with people taking all sorts of approaches to studying how to improve the quality of healthcare. I’m Jenni Burt, senior social scientist at The Healthcare Improvement Studies Institute.

Today, we are turning our attention to an increasingly high profile topic, that of citizen science, often known as people-powered research. There’s been an explosion of platforms and projects which enable anyone to be a researcher. We are going to turn to my children here for some insight into what kinds of things citizen science requires you to do and why you might get involved.

A great destination for citizen science projects is the Zooniverse platform, which I blame for my children’s obsession with spotting bears’ bottoms in the name of environmental research. It turns out that nothing is apparently more hilarious than being asked to look at camera trap photos taken in Upstate New York to report what, if any, animal was wandering past. The research project was actually aiming to look at the impact of environmental changes on the local cottontail rabbit population. But in our house, it was mostly accompanied by an ongoing commentary of, nothing, nothing, rabbit, [laugh] bear’s bottom.

So, now I’ve outlined the personal joys of involvement in citizen science, I’m delighted to welcome Michael Hornberger to the podcast. Michael, you’re Professor of Applied Dementia Research at Norwich Medical School, and also the Associate Dean of Innovation at the Faculty of Medicine and Health Sciences at the University of East Anglia.

Your research focuses on both clinical and non-clinical aspects of dementia. What’s really relevant to our discussion today is that you are well-versed in the use of citizen science approaches. You’ve created a smash smartphone game, Sea Hero Quest, which as well as being award-winning and super fun to play, gathers data on how people find their way around their environment, their spatial navigation. Millions of people have downloaded and played Sea Hero Quest, and as a result, you’ve been able to make some major insights into how people navigate spaces and what this can tell us about early disease detection in relation to dementia.

Michael, how did you first hear about citizen science?

MH: Well, you know, thank you, first of all, for having me. I think citizen science is really integral to a lot of things we do. A lot in health sciences, of course, and medical research in particular, many of us do citizen science without actually realising it. In particular, patient and participant involvement studies is a lot about citizen scienceship, that you invite lay people to basically help you with your research. It’s really on different levels that they help you; either conceptualising a research, or designing a research, running your research, or even analysing your research.

So I’ve heard of it over many years, but as you said, our Sea Hero Quest project, it took it to another level, that we actually allowed people to collect and analyse their data. So that was in addition to the PPI research we have done for many years. It really added a completely different dimension to our citizen science side.

JB: So [laugh], you’ve obviously heard my particularly helpful outline of our family’s experience of citizen science and how it works at least from the perspective of small and overly enthusiastic children. You’ve just said this, that there’s a kind of conceptualisation that citizen science falls into three different parts. So one is involving people in the development and shaping of research, the second is inviting people to help with the analysis of an existing data search, and the third is inviting people to collect data. You’ve mentioned already that Sea Hero Quest fits into the collecting data strand of citizen science. It’s a fantastic story of how you came to develop this game.

MH: Yeah, so it was, as you say, an interesting [laugh] story, for sure. What really happened, that I attended a workshop of the Wellcome Trust on gamification of science. Before that, as I said, we did citizen science on a PPI level. I was aware of the Zooniverse and other platforms, but the gamification aspect really made me think that games are an ideal way of collecting data with lay people for science. ’Cause, frankly, the collection of data in science can be sometimes a bit dry, in particular for lay people. Us, scientists, get very excited about it, but for lay people it might not be very exciting. So having it embedded into a game was really a key aspect to make it attractive for people; they want to play and therefore really collecting data.

So the actual data collection part of it is implicit within the game. Of course, we are very explicit about consenting people, that they’re happy to contribute their data. They can play the game without giving us their data. But this kind of gamification aspect added a fun aspect for people, which then they could contribute to the research while at the same time having fun playing a casual mobile game.

JB: So what actually happens in the game?

MH: Based on the research really that we’re knowing now that spatial navigation and orientation deficits in especially Alzheimer’s disease are really some of the first symptoms and sometimes happen many, many years before people actually develop memory symptoms. But we very little understand of what’s actually going on. So, can we detect it as early as possible? Because these changes are very, very subtle.

So in Sea Hero Quest, what we do, we give people this game where you have to navigate a little boat through different environments. It’s really just to collect stars, find your way through these different levels or environments, and we’re measuring how you move through those environments. Now, what people are not aware of, that behind the scenes we have basically bespokely designed the environments and the levels so that they challenge your navigation system quite…for particular factors which we know are very vulnerable to Alzheimer’s disease. So you play it, we collect the data, and therefore we can determine how healthy people navigate in these environments. Then if we compare this to somebody who might have early signs of Alzheimer’s disease or is at high risk of developing Alzheimer’s disease would perform, we would think that we can see a difference for that.

JB: It’s important to stress that this is genuinely a smartphone game. It’s downloadable from everyone’s normal app stores. It looks very much like commercial games might look. Is that right?

MH: Well, it’s not anymore actually available for download at the moment. If you want to, you can still visit the website. So if you type into Google, Sea Hero Quest, you can still find the website for it. But at the time, we started collecting the data, and of course with any study you do a data census where you stop collecting. Now we’re having a version of the game – not the exact game but a part of it – which we’ll release now for patient cohorts or research cohorts, which will be hosted by Alzheimer’s Research UK, who’s one of the main dementia charities and has been really a champion for this game as well and has been very supportive for us.

The key was that we didn’t want to create another online science experiment which can be a bit boring. No, it was really meant to be a casual mobile game which you download from the app store, as you say, or Google Play. Then you could you just play it any way you want, and you could contribute your data or you couldn’t. But most people, you know, over 80 per cent of people, decided to contribute the data because they loved the idea of playing this casual mobile game and at the same time helping science.

JB: I did mention earlier on that it was genuinely millions of [laugh] people who played. When I first heard you talk about Sea Hero Quest, I remember feeling seriously jealous when you sort of casually dropped in quite how many people you were able to get data from. Did it just take off straight away, and were you surprised?

MH: Yeah, we were very, very surprised. Frankly, we didn’t anticipate that it would become so big. We completely underestimated how big actually gaming is. Because I, myself, I have to admit, ashamedly, that I’m not really a gamer. My children are, my children love it, and many of my colleagues do, but for me it was really the concept I realised. When we launched it, we had the ambition of collecting data in 100,000 people over Western Europe and North America. We thought it was wildly over-ambitious, basically, to collect that much data, but we thought, well you have to sometimes be ambitious. But when we launched the game, basically we collected these 100,000 people in the first two days after the launch, so it was...

JB: See? So jealous [laugh].

MH: …absolutely [laugh] shocking for us. In the end, the game collected around, yeah, 4.5 million people’s data, which many people played the game also several times. So we also have multiple times when people played the game. So it’s an incredibly rich dataset, which we still…well, we have started publishing and analysing the data, but still it’s so much data there which would give us years of outcomes from the game.

JB: What kind of age range are we talking about here as well?

MH: Yeah, this was a very important aspect because of course this is targeted towards more dementia detection, or early dementia detection. For that, we should highlight that dementia is now considered to be a disease of middle age presenting at old age. So we actually want to detect changes which we can see already in people in middle age, and then we potentially prevent or treat them to delay or even potentially avoid their onset of dementia.

So we targeted the game for a whole broad range of people, but our main target was really to get between 18 years old up to 65, was our main target. But we had people even up to 95, I think, these were our oldest players who played the game. It was a fairly even distribution. We had more younger players of course – and gaming is much bigger in the younger communities – but also many, many older people, or middle-aged people love really gaming. We did a lot of user testing that was appealing to people of different age groups, and also the controls of the game, of course, are only also just about four different age groups. In particular, if you’re targeting older people, if they have more motoric issues that they can still play the game.

So there were lots of things to consider. We wanted to avoid the trap that we would just collect the classic gamer population of young white males [laugh]. Instead, we wanted to have a broad age range, and also an equal, or nearly equal, gender mix, which we did achieve in the end.

JB: So we’ve obviously been talking a lot about the fact that this is a game and this idea of a gamification approach. You mentioned wanting to not do yet another boring online science experiment, which of course is what most of us do all the time because actually it’s very difficult to think what else we could do. What have you found, do you think, the real value for the gamification approach?

MH: Yeah, there’s several levels, which surprised us again because we didn’t plan it that way. First of all, we realised that gamifying research allows you to actually create an enormous public engagement campaign without actually planning that. Because people might play the game and finding out about the disease or condition you’re researching, which really raises the profile of the whole disease which is absolutely fantastic.

I think other aspects is how they engage with the game itself was really critical as well. So very often in health research, of course we’re very much worried about compliance or attrition. So how people comply with what the research is measuring, or attrition, that people basically start taking part but then drop off. We found that the gaming approach is hugely successful in keeping people engaged and comply. Also, attrition rates were very, very little, that people really wanted to play it because the game just incentivises all the time that you want to reach this next level, or you want to continue, or you want to come back to play more. We found that a very, very successful approach.

I can see that other games… So actually what’s been launched is EndeavorRx, which is a game for ADHD children. So it’s teaching ADHD children how to control better behaviour. This has been developed over a long time. It’s a game for children. But I think this year it’s been cleared by the FDA as a treatment for ADHD. It shows you again how gaming, how powerful it can be as even a health treatment because the children love engaging with that. We found with our populations of patients that they’re extremely keen on keeping going with this. So in terms of compliance or attrition, it’s fantastic to use again in approach.

JB: Going back as well to the fact that you said four a half [laugh] million people took part, that’s not the kind of figures we would ever normally expect to see in our boring online research approach as well. So do you think gamification helps with the large dataset gathering?

MH: Absolutely. You would get the usual people who are interested in research and might find out about it, but you also get a completely different population who might have never taken part in research. That, as I said, in terms of outreach exercise is absolutely fantastic. This was the key balance in developing Sea Hero Quest, that it was a fun game to play, but at the same time a valid science experiment, in a way. That was a tricky balance, of course, because very often, as we mentioned already [laugh], science experiments are not really designed to be fun. They’re really meant to test very clear hypotheses. So this was a very, very tricky balance to strike for us, which we thought we achieved in the end. Our results so far bear art to this. But that really made it possible for so many people to engage with it, having fun with it, and just collect for us all this data, which as I said, we never anticipated to get this amount of data on so many people playing it.

JB: So obviously, from my perspective, I’m interested in things like quality of healthcare and safety of healthcare and things like that. We’re certainly exploring the potential of gamification at THIS Institute as we’re developing and rolling out our own citizen science work. I guess one of the things that maybe worries me slightly is that people somehow see that it’s disrespectful in some way, or not being serious enough in some way, to take these approaches to something that is a serious scientific subject. What’s your response to those kind of concerns?

MH: Well, I completely understand them. In a way, I know myself how it is to run on the one hand now these kind of gamified approaches, and the other side is really to run more lab-based serious kind of research. But I think one doesn’t replace the other, they really complement each other, and that’s really a key aspect, in a way. You can do citizen science and the gamification in a very serious way, and that’s what we did. We really did a lot of user testing and validation of what we did. It doesn’t mean that science has to be boring. Particularly, if you engaged outside community and lay people, just keeping people engaged makes such a huge, huge difference to that, so you can reach so many more people. Not even, I think, lay people, but also healthcare professionals. For them, it might be also fun to do something like this. As I said, you can design it that it is serious science while still at the same time having fun, but it is tricky.

JB: Sure. I mean, it also does sound quite fun from the research team perspective to be able to do [laugh] something like that. I’d certainly love to have a go.

We’re sort of moving into the space of exploring a few challenges maybe around citizen science. I mean, it sounds fantastic. Going back to our family adventures in citizen science, I know lots of listeners might be familiar with things like the RSPB Big Garden Birdwatch, where in January every year around half a million people sit for an hour and they count the birds that they see in their garden or their local park and you put your findings in online. I think RSPB have been doing it for about 40 years now, so with loads of people contributing they’ve also got this huge dataset on birds and how birds are doing over time.

Again, we like the Big Garden Birdwatch as a family. But I do remember when my youngest child was quite a bit younger, he was absolutely insistent that there were 47 blackbirds that he’d seen in an hour. Which to me rather highlights the importance of what kind of instructions and tutorials and on-boarding that you can give to people who want to contribute in this way to citizen science. So the RSPB do quite politely point out that they want to know the highest number of each bird species you can see at any one time. My youngest was quite upset when I tried to point out that maybe there was, like, one or two blackbirds who just kept on coming back and forth to the dried mealworms rather than a giant [laugh] flock of 47. So it would be really interesting to know what kind of other challenges you encountered with Sea Hero Quest.

MH: It clearly has a lot of challenges, and particularly because so little research has been done using such gamified approaches in citizen science. But having clear instructions is really vital, and in particular, people collecting data, is this data valid? How can you check whether people do what they’re supposed to be doing? For Sea Hero Quest, the data collection is implicit within the game so you don’t even notice that you’re collecting the data. Well, you’re consenting to it. But then it takes away a lot of the instructions. It’s just part of the game what you’re doing.

JB: Which goes back to this idea of one of the advantages of gamification, that it’s just implicit within the game rather than the upload and contribution of data in quite a sort of discreet way.

MH: Absolutely. Otherwise you need to provide the training, as you say. You need to provide the quality checks. How do you do that, that you let people train to criterion to some degree. Within gaming, as I said, we found there’s a very neat way of doing this.

The games companies use this term called ‘do-torial’, and that literally means that the first level, very often, in any kind of casual game you play, it really just teaches you how to interact with the character in the game or whatever you do. And we did the same. So for the first two levels we actually didn’t collect any of the data. We just gave these kind of do-torials. Then at the same time, we measured during that time how people interacted, for example, with the device, and we used that then later for normalising their performance. So if somebody was quite old and maybe a bit slow in their responses, we can actually detect that already at these beginning stages.

Then you have a lot of other kind of challenges. In particular, of course, if you do remote or online testing…and it’s very topical of course at the moment, because you lose a lot of control what people are actually doing. Not only what they’re measuring, but how they’re interacting with whatever you provide to them. You need to have some failsafe mechanisms that you can detect unusual user behaviour. We had that built in, for example, that we could detect a sudden user behaviour change quite dramatically by algorithms we built in. And with the more random impacts, we had always this classic example that it was just the dog chewing on the iPad, you know, and not anybody actually playing…

JB: [Laugh].

MH: …the game. So you need to think about that.

The other, of course, key issue which we have found after doing this project was always the criticism by the science community that you have a selection bias online. Actually any online study. I wonder if the COVID pandemic will actually change this perception ’cause so many people are now doing more online research. But the criticism is valid, that you might get people who might just take part because they want to play this game and so on, and they might not be representative of the population. We did a lot of validation after that to show that we could replicate the results in lab-based cohorts which we invited, so we think it is a very valid approach. But it’s something one needs to be acutely aware of. Who are you sampling out there in the community? Who is actually picking up whatever project you do or game you do?

JB: Certainly, I think there is a lot of concern about the kind of people who have the time, the interest, and the resources to engage in citizen science, particularly the contribution of data. I think that kind of debate around representativeness, who are you trying to reach, how can you reach them, and how can we better engage with all sections of the community, is something that seems really live at the moment.

MH: Yep, absolutely, I agree. I guess you have to see. If you’re collecting such large amounts of data like we did, it might be in the end more representative than actually the data you collect in a very small lab-based study. It’s this kind of give and take we don’t know, in a way. That’s the thing. But the screening clearly is much less in terms of who takes part in this, so you need to be quite careful what you do with it.

JB: I was going to ask about things like data security. This is research. This isn’t just a game. Eighty per cent of people were consenting to contribute their data to Sea Hero Quest, but you’ve got a pretty complex infrastructure behind that. You’ve got gaming company. The game is being used on people’s smartphones and other devices. How do you go around tackling the challenges and the needs and the questions around data security and consent and issues like that?

MH: Yes, a very, very important issue. Of course, data security is really paramount these days. For Sea Hero Quest, we took the approach that we actually collected only anonymised data because we just wanted to collect very large normative data of their navigation behaviour. So with that, the data was…even though it was saved on secure servers. But we had algorithms where the data from people’s mobile phones or the devices they played on would come onto the server. The first algorithm basically scrubbed all the information, even from the phone and where the data came from. We had therefore completely anonymised data – except for information that people provided explicitly, if they wanted to provide age, gender, where they come from, and so on and so forth – because we knew it would be potentially quite sensitive data.

If you want to do identifiable data, you need to be very, very careful that you keep this quite separately stored the identifiable information from the actual recorded data, because of course you’re creating quite a significant database for people which they could access and then use that. So this is quite a key aspect.

For us, the other interesting approaches that we wanted to collect it anonymously is that the raw data of Sea Hero Quest we made open access. So anybody basically can download the whole raw dataset if they want to. That really reassured people that it’s not just data we’re collecting for us, so for some kind of various reason [laugh], but it’s really for the common good. If anybody wanted to download the data, we’re very happy to provide it. Though, I’ll tell you, it is very, very [laugh] large, so you need to have a very large server infrastructure to actually do anything with the [laugh] data.

JB: One of the things I really did want to reflect on, and I think we’ve already touched on this a bit, is this idea of citizen science as proper science or not proper science. So it’s new, there’s a lot we don’t know methodologically, and I wondered what your experience had been. You say you’re starting to publish findings now from Sea Hero Quest. Is the scientific community a bit snooty about data collected this way? Is it sort of not taken seriously? Is that a debate that we really need to have?

MH: Yeah, I think there is a debate to be had. Yeah, we found quite a lot of resistance of publishing this data. It took us quite a while. Many people really asked us – fair enough, I would say – to provide a validation study in lab cohorts. But, of course, you can’t test 4.5 million people in the lab, so you can’t replicate the data. If you replicate the same result in 50 people or 100 people in a lab, is this representative of the bigger dataset? What does this mean exactly? I think these are really fair discussions one needs to have.

So we have found a way now though that the data really is now published and is very successful, and we’re having incredible publications coming out which we’re very excited about. The amazing thing, as we say, we opened the data, open access, so people can publish from this data whatever they want to have a look at, which is really fantastic. The patient side is only taking off really now, that we’re having so many different groups in the world now want to use this dataset as well for their patient cohorts.

In terms of publications, yes, you’re absolutely right, is this a valid way of collecting scientific data is clearly an ongoing discussion. The same, I would say, is then for the grant funding, which we have found it even harder for this because it is just too blue-sky. Now only with the publications now coming out we can see that people really believe that this is actually a potentially valid approach of collecting data.

JB: Do you think this debate is slightly different? I mean, going back to what we were saying at the beginning around the different reasons that we might consider using citizen science. The second of those is actually you have already a huge dataset and you’re asking for some help in processing or analysing it versus trying to create a dataset. Certainly, for example, within the systematic review community and coming from organisations like Cochrane and Cochrane Crowd, there’s been quite a lot of experimentation around using citizen science within systematic reviews to help for things like screening titles and abstracts. There, you can develop quite clear algorithms about how many people need to look at an abstract for you to be confident that there’s consensus about whether that abstract is relevant or not for your review, for example.

So it feels like there’s maybe a bit more control over how that contribution is made. But when we’re just asking people to contribute their data, do you think that’s partly where the nervousness comes from? We’re so used to having more control over where our data and how our data are generated?

MH: Yes, absolutely, I think it’s clear that you have much more control. You could screen much better who is doing what. You could do much better quality checks [inaudible] before sending it out, then people collect data, it’s much harder to maintain that control. But this control, of course, can be also an illusion, as we all know, because we all have screening protocols for lab studies, which might actually the screening itself creates a selection bias that actually the data in the end is not representative.

The other thing which I’m always actually really worried about is also our PPI input, to come back to that.

JB: I’m interrupting, just so for listeners who might not be familiar. We mean patient and public involvement. So this is very much getting input into the design and the setup of studies to make sure. That could also be NHS staff involvement, for example, if that’s relevant, not just patients.

MH: Yes, I think it’s very important for the PPI aspect how representative are the people who are invited to these panels. It’s very, very difficult to establish. We find that extremely hard. Other people in the room who will help you co-design these experiments, are they really representative of what you want to measure of the population you want to reach? I don’t know. The question is, is this even quantifiable? It creates all these kind of issues of engaging lay people in your research, which I think is a fantastic thing. But it’s really just as a discussion to be had, how do we generate valid data still involving lay people in our research?

JB: Having chewed over some [laugh] of the challenges within citizen science broadly as well, it would be really good to get your thoughts on where you think we’re going. What comes next, do you think? What’s kind of new areas of focus?

MH: Well, I think there’s still so much to do in terms of developing how you actually conduct this, as you say, in a valid way, and how do you engage people. I think, for me, there is a huge aspect for keeping people engaged and taking part in studies, or potentially even trials or interventions or rehabilitation. I think that has virtually not been explored. As I said, this other game which just came out on ADHD, which is FDA approved, is again a sign for this. This can be potentially a treatment option, which a few years ago would have been completely a blue-sky kind of approach.

But I think there is a huge scope. I think just involving the public, getting more public involvement, is always a good thing. To have more heads thinking about this, and getting people, or service users, who are really exposed to some kind of condition involved. I think it has been always extremely valuable to us. I think in that sense, citizen science is, I think, the entry way for universities to invite the public much more to be part of the research. That’s where I hope it would be going in the future.

JB: I’m back to this idea of the different sorts of citizen science is actually a more holistic and joined-up way of thinking, and this touches on other areas as well. We’ve had a whole set of other discussions in another podcast around patient and public involvement, how we do that, and what the challenges of doing that are. But actually the kind of joined-up thinking of the fact that it’s we’re co-producing research projects, we’re inviting people to help classify data, and we’re inviting people to contribute data at this kind of mass level. It sounds like that’s where you’re saying we might be pushing things.

MH: Absolutely, and really opening up the research laboratories to the public. I think that’s what the key aspect is. It’s science for the people from the people. That’s really what it’s all about. We, as scientists, just have the specialist knowledge to steer this, and have a much broader overview and specialist knowledge to steer and design the research that way, which I think can make a huge difference then.

JB: Certainly, again from my perspective and the kind of research that I’m concerned with, it’s not just about patients and the public. It’s also around the staff trying to reach out to everyone who’s working within the NHS. We typically tend to focus our research around the experience of nurses or doctors, but actually there’s so much knowledge and wisdom in the hospital porters and the receptionists and all the people who could be really making a difference. It seems to me that there’s a lot of potential with these kind of approaches for reaching that more diverse…

MH: Absolutely.

JB: …audience as well.

MH: Or social care as well, clearly, because we’re now broadening our tent to much more holistic again, you know, health and social care. Which, again, how does it link all up, I think is critical. I think one way is… Very often, people get invited, of course, for research from these kind of, you know, either health professionals or related to health professionals, but citizen science makes them…gives some kind of ownership of being really part of the research, and that’s really important. That they understand why is this research important, what will it help, that I contributed to that, is I think enormously powerful, in my experience, that people again are engaged and really want to make a difference. Because most people in health and social care, of course they really want to make a difference, and again citizen science is just another vehicle to actually increase their power in shaping the future of research and eventually their services.

JB: So you mentioned you’re obviously developing Sea Hero Quest in more patient platforms. Having learnt everything you’ve learnt, would you start again from scratch and do a whole other smartphone app game to give new insights to another aspect of your research?

MH: Yeah, I’d love that. I think I’m all up for that. Because, of course, we’ve learnt so much we could avoid so many of the naïve mistakes we made the first time round. In hindsight, you’re always far cleverer, obviously, as always.

I have to say, I’m involved in some other kind of studies which have taken on this approach and also some companies who are exploring these kind of approaches. I’m very happy with that because I think, again, if more people try this approach, the better. I, for myself, as I said, we are only now reaching that we can publish it…well, we have started publishing the data and having also the patient site going. After that, I definitely will look for more funding to do future research in this because, as I said, it just complements our lab research so beautifully, and gives you a very different angle, and makes this whole public engagement a fantastic concept which just empowers everybody and makes a real difference in the future.

JB: We’ve covered, as per normal [laugh], so much ground around citizen science. I think it’s really clear from the example of Sea Hero Quest quite what potential there is in citizen science approaches for gathering amazingly useful data directly from people in a really fun way, so that kind of opportunity for engagement. I also think gamification really does sound like fun on both sides, not just for the people [laugh] who are using the game, but it would be so much fun as a researcher, I think. But it’s also really clear from the challenges that you’ve encountered that this is new territory, and there is a lot of methodological work and development that’s still needed to answer questions around validity of data, particularly around representation and making sure that we’re really working out how to engage with all sections of the relevant communities. So a lot still to be done.

Thank you so much, Michael Hornberger.

MH: My pleasure, I hope it’s helpful, and helpful to your listeners. And yeah, please go out and do some citizen science.

JB: I’m Jenni Burt, and you’ve been listening to THIS.

**End of transcript**