Calling DIBS

Episode 56: "Comparing Lab, Survey, & Field Research Designs"

with Vince Hopkins, Assistant Professor of Public Policy at the University of Saskatchewan

Vince Hopkins returns to the podcast to chat about the continuum of research design from lab studies to field studies. We discuss the opportunities and challenges of different research designs as well as how to deal with different constraints prospectively in design or retrospectively in analysis. We also discuss how field studies provide a wealth of learning for all involved, no matter what the results are or whether a trial was even run.

Transcript:

KIRSTIN APPELT, HOST: Welcome to this edition of Calling DIBS. I'm your host, Kirstin Appelt, Research Director with UBC Decision Insights for Business and Society or DIBS for short. Today we're calling DIBS on Vince Hopkins.

Vince is Assistant Professor at the University of Saskatchewan's Johnson Shoyama Graduate School of Public Policy and one of our wonderful instructors in the Advanced Professional Certificate in Behavioural Insights. I first connected with Vince back when he was Senior Behavioural Scientist at the BC Behavioural Insights Group, and we rarely get to talk shop these days, so I'm really excited for a chance to reconnect and nerd out on research, design, and all things quantitative. So welcome back to the podcast, Vince.

VINCE HOPKINS, GUEST: Thank you, Kirstin.

APPELT: Well, you have changed roles and even provinces since you were last on the podcast. I thought maybe it would be good to give us a little update as we get going here. So where has your BI journey taken you since your time at BC BIG?

HOPKINS: Thank you. Yeah. So, two years ago, I accepted a job here at the University of Saskatchewan. As you mentioned, I'm in the School of Public Policy. And actually, it's been a great journey because the school integrates a lot of people in political science, but also people in economics and behavioural insights, and especially applied behavioural science borrows a lot from both of those disciplines. So, I am surrounded by political scientists; we get to think about things like ethics and power and how people think about government and how government should deliver optimal services. I also work with economists who think about, certainly from a behavioural perspective, consumer behaviour, how firms behave.

It's really in some ways pushed me and helped me think about behaviour change more ambitiously, more seriously. But I still get to do the same stuff that I got to do at BC BIG. So, my BI journey has been pretty great the last two years.

APPELT: Sounds like so many interesting conversations are being had there. I love that. Well, so last time you were on the podcast, we were talking about data, research design, and ethics, and I thought it'd be great this time to focus more on the research design piece and specifically to explore lab and field studies, because I think that's an area where there's a lot of interest, yet it is sometimes hard to find good readings or materials to teach that. And I think a lot of it comes with experience. So maybe we can just start with the basics and you can explain just upfront, what do we mean when we say lab studies and field studies and what's that distinction?

HOPKINS: Yeah. So yeah, so you mentioned lab studies and field studies. I probably also mentioned survey studies because we're seeing more and more companies, firms, governments using survey experiments to understand behaviour change. And so, I kind of want I to talk about all three I mean, lab experiments in some ways, people who've read books like, you know, “Nudge”, but for sure “Thinking Fast, Thinking Slow” have been exposed to lab experiments because a lot of that early research, even as far back as the fifties, certainly in the seventies and eighties, a lot of that came out of lab experiments. And so, a lab experiment takes place in a controlled environment. And usually that's a physical setting, like a laboratory, often on a university campus. And the idea here is to have sort of a precise control over what variables are being tested. And so, we have a high degree of control over the actual experiment.

You can sort of think of survey experiments as a move away from the control towards realism. And in fact, some people think about experiments on a continuum of realism. So, a survey experiment is kind of like a lab experiment. People, you know, sign up for an experiment. They're randomly exposed to one experience or another. But in the lab experiment, you know, it's in the bunker of some psychology building at a university campus, with like a flickering fluorescent light. In the survey experiment, it's collecting data over -- it could be over the phone, but these days, mostly online. And these are people that are doing these online questionnaires. So, this gets a little bit closer towards realism.

And then, you know, as you move from lab to survey, you find field experiments. Which I mean, some people say they take place in a natural setting. I don't really know what a natural setting is, but the idea that it's in a setting that is natural or sort of organic to the research questions. Could be in a community or maybe there's a specific geographical location where the research has to take place.

When you see studies done in the lab, the survey, the field, researchers have made that choice based on all kinds of factors. And, you know, we can talk about some of those factors, why you might think about a lab or survey or field experiment, but at the highest level I think about these three, sort of on a continuum of realism and control.

APPELT: Yeah. And I'd love to peel apart those a little bit. I think like you said, it's great to think of them on a continuum. So, I'm wondering if we can get into the similarities and differences. And so, one of the things you mentioned with lab studies, the prototypical flickering light, university psychology building basement, for example, one of the critiques of that is that it has a quote unquote “WEIRD population” where it's perhaps sampling from undergraduate freshmen, sophomores, first year, second years. And so, then if we move into survey studies, you're getting perhaps more typical populations that actually reflect the world outside of a college campus and field studies, perhaps even more so.

Do you want to elaborate on that? Or whether other similarities and differences are there?

HOPKINS: Yeah, I mean, that's certainly true. And, you know, I guess I tend to think of the sample question in terms of control, like, you're right, a lot of those cities are, you know, WEIRD: Western, industrialized, democratic, etc. But, you know, from another perspective, it does mean that they've been carefully selected. Like, the researcher generally has a very good sense of who these people are, undergraduate students in psychology or economics. And to some extent, that has, you know, maybe helped replicability because if you're doing one experiment with one group of 20 something year-old psych students and you want to replicate it, you can sort of rely on, well, the samples are kind of similar. So, the everything is, yes, location specific, attribute specific.

You mentioned the WEIRD acronym. So, a bit smaller and in some contexts, that's great. In another context, that poses a trade-off. I think the most obvious trade-off is to the survey experiment, and I think that's where you were sort of getting that where we see, yeah, now we can have much larger sample sizes and you're seeing this more and more that like if you open up on one of the top journals like *Nature* or *Science* and you see a survey, it's like a survey of 70,000 people and they've got these enormous sample sizes. You know, I think in that sense, we're increasing the diversity and the representativeness of our samples.

But it's not without trade-offs here as well, right? It's only certain types of people that do these surveys, especially online surveys, right? I work a lot with polling firms. If you go through a polling firm, they have these in-house panels, I give the money to the firm. The firm gives the money to the person on the panel. And that person, the panel is probably maybe a bit systematically different than people who aren't in the panel. And so, we here again, we see this issue of the representativeness of the sample. I mean, you give a bit of control here because you don't in the lab experiment, you know, the setting in which all the answers are mean, all the questions are being answered. You don't have that in a survey, so you lose some control.

Some people call this “mundane realism” and Diana Mutz -- who's a political scientist, emphasizes how mundane survey experiments can be. These are often people who are in their homes, on their desktop, computer or laptop doing academic surveys or, you know, consumer-facing surveys. And so mundane realism is sort of how much does this experiment mirror the real world we're trying to extrapolate to, we're trying to understand.

And so here again, we're sort of sensitive to sampling. People self-select in or they only selectively respond to certain types of surveys. Similarly, we see in both lab and survey experiments is this sensitivity to measurement error, right? Like your understanding, your knowledge is only as good as your survey questions. “Do our survey questions measure what we think they measure?”. Measurement error also is a factor that's common to both lab and survey experiments. Yeah, I think lab and survey experiments share a lot because they often, not always but often, have sort of a root in survey questions.

APPELT: Yeah, absolutely. And something else that I think comes up there when you're talking about the idea of mundane realism and the professional participants who this is like a job for them, they're getting sometimes very little, sometimes more generous compensation, but there's had to be a move to actually make sure that the same participant, if you're doing a series of studies and they build on each other and they're using the same questions, you have to make sure the same participant hasn't done all of your studies, for example. So, I think that's something for folks who aren't as experienced in this world may not realize that that is something that the field has had to contend with, is the idea of people who are repeat participants and how that can affect your research results.

HOPKINS: Yeah, totally. In fact, there is some great research done in political science on the effect of COVID on survey responses, right? Suddenly, you had all these people at home able to do surveys on their part time and make a little bit of extra money. And that systematically changed the composition of who was doing those surveys. And there was a big sort of question, “If these people are different, how are they different?” And one way they might be different is they might pay less attention. Sometimes they pay more attention.

But the idea is that attention, which is really what we're after in most applied settings, attention is a moving target and some samples might get you more attention. Others might get you less attention. So yeah, you're totally right. Who these people are is not just about sampling. It might also get at sort of the inferences and the noise in the data that we're collecting.

APPELT: Yeah. And there's even been interesting work on, for example, MTurk, which is a very popular platform where certain known and replicated results not getting into the issue of some results that don't replicate, but things that replicate really well were no longer replicating because if you've seen the Linda problem from thinking fast and slow 14 times, on the 15th time, you're probably going to get it right. Not because you necessarily get the reasoning, but just you read it before you know the answer. So, there's a lot of interesting anecdotes from that side of things.

HOPKINS: Yeah, this is a big question in applied behavioural science research, whether it's social psych econ or political science, sociology, you know, what have you, is there's no sort of perfect free sample. Everything has a cost. You're always, as a researcher, trying to balance "How much money do I have? How much time do I need from people? How can I get the most attention, the best responses from the right kind of respondent for my research?".

In my experience, each project is a bit different and I work a lot with polling firms, as I mentioned, and I like that because I get to sort of talk to them in advance about my expectations about drawing an attentive sample, and if you use things like MTurk, that's harder because you're negotiating directly with the survey respondent, so it can be nice for polling firms. You can save a bit of money if you find a good firm that's willing to let you, for example, include what we call an attention check, you know -- a question that we know there's a right answer to it. And if you give the wrong answer, we sort of kick you out of the survey or we invite you to leave the survey, and that person might not get paid as much. That can be harder to do with MTurk, if you're recruiting your own sample. It can be easier if you're working with a polling firm. But it all depends on the firm. It all depends on the respondent.

This is, I think, part of the research skill that certainly grad students get a lot of nowadays. It's probably changed a lot from even 20 years ago, 15 years ago, how to recruit sample as much as sort of methodological intensive bit of training.

APPELT: Yeah, there is a bit of an ongoing race to stay ahead of things like bots and other things. When I was doing some of my graduate work, there were known attention check questions like "Have you previously died from a heart attack?" People would say, yes, well, you're probably not paying attention because if you died, I don't think you're taking the survey. But then over time, you know, that stopped working because everyone knew to look for that question because there are message boards. So, it is interesting how it evolves over time, and attention checks and everything have to keep pace.

But I thought maybe because I think for our audience, they're more often on the field study side, so maybe we should shift a little towards the field studies. So, we've talked how lab and survey kind of maybe are closer on the continuum. Are there similarities, differences with field studies you wanted to draw distinctions around?

HOPKINS: It's first, I think, helpful to think about what is common to all three types of studies, right? And for me, like, the thing that I live and breathe is the randomization component, and that can go wrong in lab experiments, it can go wrong in survey experiments, and go wrong in field experiments. And at the heart of all three of these designs, we can think about all the reasons are different about the samples and stuff like that. But the thing that unites them all, the thing that excites researchers like me, I think probably a lot of your listeners is, is the randomization. And in fact, we'll talk about that I think later probably about field experiments.

Randomization is especially challenging with field experiments for a whole host of reasons, but some of the things that are great about field experiments, sometimes it can mean you have a larger sample size, sometimes even an incredibly large sample size. I think about a lot of, you know, my wife works in tech, and a lot of tech firms when they do A/B tests, which in their context might look a lot like field experiments, they might have, you know, a sample size in the millions. And, you know, that's generally wonderful. We generally want more sample than less, although sometimes we might have too much sample, right? And we've wasted people's times in an experiment.

In field experiments, one of the trade-offs here -- yeah, you get this bigger sample size, you can claim, I think more plausibly, that you have a more realistic setting, that your inferences and your conclusions are more applicable to real-world situations. But one of the big trade-offs here is like control, right? And so that's, I think, one reason why lab experiments and field experiments are thought of as two ends of a continuum, gives you just, you lose a lot of control. I love that. That's my favourite part about field experiments is it's so messy and everything can go wrong. And I mean, I just love that. But it's a huge, it creates an enormous challenge on the researcher, on the partner who's actually running the experiment because a lot of these are generally partnership-based. It introduces questions of data sharing, of costs and all these things that, you know, are generally much easier to address in lab and survey experiments. I would say also, I think there's a greater variety of, I think creativity is a real premium in field experiments. And there's a lot of really great cases where people combine, for example, surveys with field experiments, right?

So, like David Brockman and Josh Kalla have this really, really cool experiment a few years ago in Miami where they send out these canvassers who go door knocking. And some of those canvassers were people who identified as transgender, others did not identify as transgender. I think they went out to like 500 something people and they had these ten-minute interviews, right? So, each canvasser is assigned to a house. They knock, they have a ten-minute conversation, and the randomization comes in that some households were randomly assigned to have a conversation about recycling. It's a sort of innocuous conversation, but another group of households were randomly assigned to talk about transgender discrimination, a much more serious, much more intense conversation. And before and after the interview, each canvasser had a survey with them, and they had the person fill out the survey like in real-time before the conversation and after the conversation. And then they even sent a follow-up survey for the next few months. And, you know, this really cool combination of a field experiment, door-to-door knocking with real people, combined with a survey experiment means that they could actually make inferences about conversations, about interviews, about conversations, about really hot-button issues. And what they found was the door-to-door canvassing and this ten-minute conversation strongly reduced prejudice towards transgender individuals. That's interesting I think because normally like discrimination can be very difficult to observe. Prejudice can be very difficult to observe.

I love this example because it's fully a field experiment. It's out there in Miami, in the real world, but they use a survey to measure something, what we might think of as latent very hard to pick up on. So, yeah, I think these experiments are really exciting and really rich, but they're also sometimes hard to characterize because they can accommodate these various elements of it. And people like John List, who was the Chief Economist for Walmart have done really great work trying to categorize well what types of field experiments are out there.

APPELT: Yeah, absolutely. There's such a variety there, which is part of what makes it exciting, but part of what makes it difficult. And I think for people who are coming out of a PhD program where maybe their first few years they were doing mostly lab or online survey work where it was: “Oh, I'm going to run a survey”, you run it and it's done same day. Then you go into a field project and it might take months just to get the permission. That can be a world of difference. But like you said, they can be very complementary because of their different strengths and they can be nice. And in some projects, you're able to actually do both things, like you said.

HOPKINS: Let's say you're in a project, and the partner doesn't usually just say like, "Hey, we're going to have an issue with this". Like, how do you solicit the information to help you identify where challenges might be? Are there certain questions you tend to ask or certain things you feel out? How do you approach that? There is a method to field experiments that is really rich. And because the technique is not that new, I think computational costs, increase in acceptance of randomization as the gold standard of evaluation, mean that we simply know a lot more about experiments now than we did even 25 years ago.

I mean, some of the big problems that we know show up and that everyone will see at some point are things like attrition, right? So, like you want to see if you can increase sales of delicious gelato, but you have sort of systematically missing data on gelato sales. You know that people go into the gelato store, but you don't know if they actually like, did they buy gelato or not? You're missing data systematically and you might have this sort of this missingness problem that can create all kinds of issues, especially if it's imbalance between control and treatment groups.

Sort of a related problem is that, it's a really big one in field experiments, is noncompliance. And this can get sort of technical and mathematical, but also has like, I think a pretty intuitive sense in government, like the way I like to think of it is governments send a lot of letters to people, right? I did a tour once of British Columbia's mailing facility and they have like a full-on post office. It's incredible. Governments send a lot of letters to people and governments generally want to know the effect of sending letters on behaviour.

Academics, though we generally want to know the effect of people reading those letters and those are actually not the same thing, right? It's really easy to measure if we sent out letters, it's really hard to measure if someone read the letter and you can think of this as reading the letter as a kind of compliance. We assigned this household to treatment, which means we sent them a letter, but we don't know if they actually read it. We don't know if they complied with treatment. And so, noncompliance is another one that comes up. And sort of means for me as a researcher, I go in and I say, "Okay, what do I care about or do I care about the average treatment effect among the treated? Or do I care about sort of what we call the intent to treat effect?" In my experience, governments care a lot more about intent to treat than the average treatment effect.

Spillover effects are another issue that comes up a lot. It's what gives birth to the sort of clustered randomized controlled trials, which I think a lot of people like, they read about them and they want to do them and they're like enormously difficult and super hard to do with sufficient statistical power. So, spillover effects are something that like there are design-based ways to solve it. That is, you can design a cluster-based trial or there are like analytical ways to address it. You can sort of use math to estimate the extent of spillover. These are the big problems, right? Like attrition, noncompliance, and spillover are some of the biggest problems, the biggest threats. And they generally have, as I said, a design-based solution or an analysis-based solution.

And I think this is where training and programs like UBC's Advanced Professional Certificate in Behavioural Insights are so powerful because without the training, you might lead an enormous field experiment and find out after the fact you have differential attrition. You have more people dropped out of your study in treatment than control. Oh, no, what do you do now? And so, I think training is especially important in field experiments because of the costs of data collection.

APPELT: Yeah, absolutely. And it also relatedly, going to the solution design, part of it too is you may not realize that there are other systematic differences between conditions. So, let's say you were trying to do a controlled condition and a BI solution condition, and you thought the difference was that you added a photo, but you also changed the text, and the text ended up being a lot shorter, so it was faster to read. So, I think the more training you have, the more you're attuned to what kinds of things to look out for.

And across these things, like you said, whether it's attrition or whether it's whether the conditions are, how the conditions vary and do they vary in the ways we want them to vary? And also, are they similar in the ways we want them to be similar?

HOPKINS: I think yeah, that's such a great example. But I think the truth with field experiments is the only way you learn is by failing. I think that just like it's like built into the technique. In fact, there's a great book called *Failing in the Field*. It's a book just of failures about that's especially about working with partners. But failure is a big part of field experiments.

And, so that's why I think it also means going into it to build like a really-- Sometimes like to say that the real experiments are the friendships you make along the way, like the partnerships and the relationships that you have with your partners have to or ideally last beyond the one trial because things might go wrong in this trial. If you have a great relationship, you can do it better again next time. And so, building in these relationships to sustain the learning, that's been a big part of my own approach as I've learned over the years.

APPELT: Absolutely. And I think there's also that hidden piece of learning that happens too is as you fail together on an experiment, you're both learning how to better do the experiment. So not even just the, you know, like, “Oh, next time we'll watch out for attrition” but “Hey, in our organization, it turns out we really need Jean from Accounting to be involved because she's the person who controls X that we need”, or “We learned that our email-sending software is buggy and it's going to send things in HTML format completely illegible”. Just like those learnings that you get over time as well as the things of like how to make sure the designs are similar or different, all of that other piece.

HOPKINS: Yeah, that's awesome. One of the big lessons that I had when I went from -- I mean, I've always gone in and out of academia into government most recently I went from, during my PhD, into government was: I had to shift what I thought a successful field study looked like. And I went in thinking, “Oh, a successful field study is one that has like high scientific validity, is one that has strong theoretical contributions”. And I came away being like “Was data collected? Did we randomize something?”.

And that for me was like, that was a new, you know, measure of success. So, I had projects in government where the partner was so keen to experiment, and I, I wanted to nourish that and support it. But the conditions were not right for a really top-notch experiment. But we did it anyways because I just wanted to show them we can build this culture of experimentation and we'll get to stuff like validity and theory. We'll get to that later. So, I think that's really well said that sort of I think you said "failing together." That's great.

APPELT: Love that. Well, I think we've done a good job of kind of mapping out some of the challenges once you know where those challenges are and whether it's attrition or whether it's design or whether it's just something as simple as, “Hey, our email system can't send out different versions of an email”. How do you decide where are places that you can make design compromises? Where do you hold the line? Where do you say, “No, this is just going to be a misuse of resources because it's just gone so far that it's not testing anything we're testing”.

HOPKINS: When I meet a new BI practitioner, I love to ask them like, “Have you ever struggled to do an experiment?”. Because I know the answer, right? And I know they're going to say “Yes”. And then I get to be like “Why do you think it's hard?”. And I've asked that question of like everyone I've ever met who works in BI, and we all have our own sort of folk theories or pet theories as to why it is.

Dilip Soman at the University of Toronto had a, I think is for me a really compelling framework, where he talks about the costs of experimentation. And I love that because some people are like, “Oh, it's hard to do experiments because people don't care about data, people don't care about evidence”. And the argument is much more generous is that people actually might love to do an experiment, but they're really costly and doesn't have to be financially-- it could just be in time, in attention, in hassle. And that certainly dovetails with my own experience.

A lot of my work, a lot of how I sort of like map out when to compromise, when to hold the line, where to make a no-go is first trying to figure out, you know, “Where are the costs for my partners and what can I do to help them reduce the cost?”. I might start out by asking things like, you know, like, “What keeps you up at night? What is your organization trying to change? What are your priorities right now?”. And I might slowly talk about data and be like, “Well, how do you currently measure that?”, or like, you know, “Do you have any variables or indicators that you use?”, and then maybe talk about priority populations, target populations, samples I might then get at like intent to treat versus average treatment effects and how reasonable that that stuff is. I might think about the approval process.

And so, I get these questions that I try to make very friendly and they serve the purpose of building the relationship, but they also help me understand what's the biggest threat to the project here? How can I help bring down the costs of experimentation and do I think I'm going to be successful?

And that tends to be how I make the sort of decision to either compromise or make a no-go decision. So, if a study's really, really important, you know, like, the project partner is like “That affects this really important population” and it's like it's critical for them, I'm like, “Okay, well then we really got to do it well”. Like there might be an ethical imperative in that, in that sense, to do this properly and to take our time at it. Maybe there's like a practical feasibility thing. If I can make a simple compromise on my end, you know, like, have one of my research assistants do a bunch of manual labour to simplify the costs for them, then that's easy for me to do.

Sometimes there's like an ethical dimension where those people might want to make a compromise on the ethics. And that is just like a nonstarter, especially now that I'm in a university with a research ethics board that I can talk to and work through these questions together. The ethics is really important, and I think one of the values added that the university researchers bring is a real serious attention to ethics. I think it's actually like something that we can contribute. And then for me, now that I'm an academic, is the scientific validity of it. Like, “If we do this, is it going to weaken the contribution of this project and how much my willing to weaken it before I say this really isn't worth collectively all of our time” and then I walk away from it.

That decision of, you know, compromising and holding the line -- calling it no go, it's really hard. And I generally I try to salvage something. And sometimes it means I turn to survey experiments like, “Oh, we can't run a field experiment, how about I use one of my research budgets to pay for a survey experiment on your behalf”. And then we were randomizing something and we're collecting data on something. They're thinking critically about behaviour change and theories of change. For me, that's a win-win. The relationship is being strengthened. I love it. I'm very happy with that kind of outcome. Yeah, absolutely.

APPELT: And I think it's also true, too, that it's not necessarily that there is a decision point where it's compromised, hold the line where to make a no-go. It's often like a series of those over the course of the design, like, “Okay, we've decided that it's going ahead”. But now you're saying that you can only send two types of emails instead of three types. So, it's often like a continued puzzle that you're continuing to work through together.

So, is there any of the smaller decisions that come up around either like randomization or what is actually going to be sent out? Any advice for the smaller decision points that aren't just like a global no-go button on this feature?

HOPKINS: Yeah. I mean, my advice to a new BI practitioner is to watch out for it can be sort of like boiling pasta, like it starts off really slow and it gets hotter and hotter and hotter. And by the end, you're doing a project you have no intention of doing. I kind of want to say it's like a wet noodle. I want to push that metaphor, that analogy. But the point is that things that might not seem like a big red flag at the start, like you send them an email and it takes them seven days to write back. Maybe like, "Oh, they're really busy". You agree on a certain type of research question and then a month later they want to do something completely different, you might be like, "Oh well, they're still figuring out what they want to do" and they tell you they can get a certain kind of data set and you find out 90 days into it, no that data, that's not possible. You're like, "Well, at least they're learning".

Certainly, with the wisdom, you realize, "Oh, this project is not a good use of my time". And the more time I spend on this project, the less time I have on another important project. So, my advice to a beginner practitioner is to be really attentive to like a series of red flags. They might just not be that into you, and they might just not be in a state of readiness to do this project. And it's actually a service to them to walk away and allocate your time to someone who deserves you.

APPELT: Someone is into you, even if that partner organization is not that into you.

HOPKINS: Exactly. Exactly.

APPELT: Yeah. And then I think also, too, sometimes when you start to realize the other benefits that come out of the project, like you said, with having different definitions of success and failure, not running an RCT doesn't mean the project's a failure. Usually, a lot of lessons came out, which maybe you're able to collect a different type of data, or maybe you've just increased their readiness to do a project in the future. So, a year from now, a couple of years from now, they come back, and they have that mechanism to collect data that they didn't have before. So, it's very rare that the project is an actual failure. It's just that it produced different insights than you thought it would produce.

HOPKINS: Totally. And all the other things that we learn in BI training can also add value. Basic data analysis. Sometimes I've had projects where like it's not like the partner lacked data analysis capacity, they didn't have the time. They sent me the data. I help them understand what's going on. That's a value add. And some of my best projects were like that where we tried to do an experiment. It didn't work. We spent nine months hitting our heads against the wall, but we added value. We built a relationship that a year later they came back and whoa, now we can suddenly do a great big field experiment. So totally that's why I yeah, I do joke that the experiment, the real experiments are the friendships we make along the way. I also believe it though.

APPELT: It sounds like the magic *My Little Pony* tagline. "Friendships are magic". And also, there are experiments. Well, I do notice that we are running low on time, so I'll transition to our famous last question, which is, "Do you have a message for our new BI practitioners in training, whether it's specific to designing research or something else?"

HOPKINS: I mean, I'd say trust your gut. Like a lot of the stuff on research design are things you can learn through coursework, through reading on YouTube. You know, you can learn how to deal statistically with differential attrition and stuff like that as best you can. But your real best asset is your gut. Your gut about a project is just going in a direction that you think is constructive, positive, and a good use of your time and the same things that make most BI practitioners so excited and enthused and thrilled with the technique is the same intuition that's going to help them avoid projects that are not the best use of their time. So, my advice is to trust your gut when building a relationship and when scoping a project.

APPELT: That is very smart. Well, as expected, it's been really fun to pick your brain on research design and hear about some of the new things you're up to. And I continue to be an eager observer of all the work that you do. So thank you for taking the time to talk with us today, Vince.

HOPKINS: Thank you, Kirstin. I'm a huge fan of Calling DIBS and really appreciate being invited back on.

APPELT: Well, thank you. And thanks to my listeners for joining another episode of Calling DIBS.