Calling DIBS

Episode 20: "Using BI to Scale and Report on Projects Effectively"

with Sasha Tregebov, Director of BIT Canada

Sasha Tregebov comes back on the podcast to talk about what happens once the results of a BI project are in. Sasha shares recommendations for how to increase the likelihood that learnings from a BI project are implemented, whether that means scaling up a given solution to a wider population, replicating the solution in a similar context, or helping the client use their experiences to inform the next project they tackle. We also discuss strategies and tips for communicating project results effectively to different audiences.

Transcript:

KIRSTIN APPELT, HOST: Welcome to this edition of Calling DIBS. I'm your host, Kirstin Appelt, the Research Director with UBC Decision Insights for Business and Society, or DIBS for short. Today, we're calling DIBS on Sasha Tregebov, Director of BIT Canada. And I'm thrilled we were able to lure Sasha back to the pod.

Last time we talked to Sasha about problem discovery, which is basically the front end of BI projects. And today we're asking Sasha about the other end of BI projects - reporting and scaling. So, it's a BIT of a Sasha sandwich, if you'll excuse the terrible wordplay. As always, I'm looking forward to an interesting conversation with Sasha, since he's someone who I admire for thinking deeply and carefully about BI. So welcome back to the podcast, Sasha.

SASHA TREGEBOV, GUEST: Thank you very much. Thanks for having me.

APPELT: And can you just refresh our listeners a bit on your role?

TREGEBOV: Yes. So, I am the Head of the Canadian branch of the Behavioural Insights Team. So that means that I get to oversee all of the wonderful applied behavioural science research that we're doing with organizations across the country.

APPELT: Fantastic. And let's just jump right into business. On scaling, there are various different, somewhat overlapping methods for evaluating whether to scale, cost benefit analysis, ROI calculations, statistical significance, practical or managerial significance, effect size. How does BIT make recommendations about scaling?

TREGEBOV: And I would suggest that, you know, as a starting point, it's even more complex than that, sometimes we think about what are the right metrics to understand whether or not to scale the intervention itself, to have the new approach that we've tested become the new normal. But there are also questions about how do you scale up the methodology of behavioural insights and causal evaluation within an organization. How do you scale up the insights that you've learned from a successful trial to other areas of focus? With regard to the sort of core question of whether or not to scale the intervention itself, I think that a sort of return on investment or cost benefit analysis is a good place to start. So long as it's sensitive to and careful about the significance and sort of confidence intervals of the result.

Generally, a cost benefit analysis would use a sort of point estimate of the treatment effect as a starting point, and that can be dangerous where you have wide confidence intervals. In other words, where you don't have a precise estimate of your treatment effect. But, you know, not only is that already a complex starting point for just this one dimension of scaling, what I've learned in doing this work with a lot of organizations is that even when you can do that cost benefit analysis well, there may be other intangible reasons to consider scaling, if it doesn't sort of fully pass a cost benefit analysis.

So, for example, we've done work with service delivery focused organizations, organizations delivering social services. And we've had situations where we didn't have, for example, a statistically significant or economically significant impact on outcomes for the organization's clients, the people that they're there to serve. But the staff at the organization thought the new approach was way more motivating to them. They felt like they were able to do their job better. They were sort of happier in the workplace. And even though that wasn't the sort of primary or secondary outcome specified in our analysis, and it would be hard to integrate that, hard, but not impossible, to integrate that into an ROI type calculation, these considerations are still really important and we're thinking about.

APPELT: I really love that. I love the point about the larger scaling of methods and learning, so we don't want to just scale the intervention, but learning how to do that process, I think that's a really good point. And then I love the idea of the somewhat intangible parts of an intervention and the idea that taking into account the experience for it, for the folks, whether it's the front service workers, whether it's the data team or whoever. I think that's a really, really valuable addition.

A lot of times when we do a project, we're really excited about the results. And maybe there's reasons it doesn't get scaled. What are some of the challenges of scaling? Why does it sometimes fail?

TREGEBOV: So, I think sometimes there are failures to scale successful interventions because scaling was not really considered from the outset of the project. The project was framed as a sort of research or innovation project in particular, and the sort of pathway to scale was not identified and appropriately secured from the outset. Something that we like to do as an organization is from the earliest conversations that we're having with an implementation partner, asking the question, “What will you do if this intervention is successful? Do you have the approvals? Do you have the resources? Do you have the stakeholder acceptance that you need for scaling?”

Because I mean, some organizations are simply interested in learning, and in the research. And of course, that's wonderful. But, you know, we think for most of our projects, if there is not a clear vision of scale, we may not be asking and answering the right research question. I think centering on scale as a way of evaluating your research question, of evaluating the impact of your project at large is an important and surprisingly challenging place to start. I think even if you get all of that right, you can still have failures to scale, and there's a lot of work that's been done, if you think about this at a sort of macro level of local intervention scaled up nationally or, you know, there are sort of questions like that and quite interesting technical answers, especially from the sort of economics literature, I'm not very conversant in that.

But from my own experience, I think another part of the problem is that scaling in itself is a behavioural problem, right, too. If you think about fundamentally scaling, it's about how do we take in evidence-informed practice and have that practice be implemented by organizations and their stakeholders? That sounds like a behavioural insights project. And it is, right? What are the cognitive, psychological, behavioural barriers to the practice being scaled? And I think you can really think about scaling as a full sort of project in itself that demands that level of thoughtfulness and attention that you would apply to your initial research, if that makes sense.

APPELT: Absolutely. I think that kind of pulls on or enhances some of the conversations we've been having around using BI in what we do, not just doing a BI project, but making sure we're using BI principles in how we communicate to the client and how we try to do scaling. I think that's absolutely right. We're asking behaviour change to get the scale, so we need to use behavioural insights in that as well.

And I also love your point about really scaling is the end, but it should also be the beginning. Scaling should be a consideration even at the very beginning. And I think that's maybe often in our eagerness to get going on a project, we don't spend enough time on scaling upfront.

TREGEBOV: That's been my experience. And I think if I could add to that something that you said reminded me, sometimes you can do that right up front and then you lose the thread. And in particular, in my experience, you can lose the thread when you start working through all of the sort of feasibility and implementation considerations related to your trial.

What can happen is you can have some like research drift where your research question drifts as you try to identify a feasible way to implement and evaluate your intervention. And you end up answering a different research question than you had started. And that's challenging for a whole bunch of reasons. But it's particularly an issue for scaling when you get done with that trial.

APPELT: Absolutely. One of the principles or techniques I teach when I teach market research is the idea of backward research, which is you run it mentally forward where you know, this is each step. But then you also need to run it backwards. So, “Where do we want to end up?”. And then at each point during the research process, say, like, “If this is where we need to be to scale, are we collecting the data for that? Are the process is going to work.” So always doing that. Each time you get to a new phase, make sure you do that forward, backward thinking.

TREGEBOV: That's really interesting.

APPELT: You've brought up already some great tips on how to promote and encourage successful scaling. Are there any other tips, maybe things once, you know, if you've missed that opportunity at the front end, but there's possibilities, is there anything you can do at the back end to encourage scaling?

TREGEBOV: Yeah, I think it's valuable to have a broad concept of “scaling”. Right, so are there other organizations that could implement this approach? Again, are there opportunities to apply these methods, you know, these lenses to other substantively similar problems that the organization or opportunities that the organization is confronting. I'll give you a short example.

One of the great trials that we've run at Canada this year was about increasing access to free tax filing services for low-income people in the Halton region of Ontario. It's just this amazing trial, I think partly because it just sort of put dollars in the pocketbooks of families that could really make good use of those resources. But the intervention that we used, which was around communicating the service or uptake of the service through an enhanced active choice intervention, in other words, really trying out for people the implications and prompting a sort of reflective decision on “Do I want this free tax filing service or not”, with the client we recognize that the barrier that we were overcoming was the same for a wide range of their services.

And in fact, they could scale this new way of communicating service offers, in terms of using enhanced active choice across a range of the services that they provide, and I think, you know, that's the sort of dimension of scaling that is worth thinking about.

APPELT: Absolutely. And that brings me to another question, which is when we're scaling, we also talk about how we have to have the evidence base for scaling. So how do we determine the limits of results? How far can we scale the results of our project before we're potentially exceeding the bounds of the evidence base and we're risking overgeneralization?

TREGEBOV: Yeah, it's an extremely challenging question, I think, to answer in a formulaic way. I think it's one of those questions and it's always a bit disappointing when you need to say this, but I think the reality is sometimes you do, it's like, it's a question of judgment.

And I think what that means is if you can bring together a range of experienced and knowledgeable practitioners with sort of different pockets of expertise, but in this case, in particular, people who really understand the theory of change around, for example, enhanced active choice communications and framing, and then pair that expertise of knowledge of the theory of change with an ability to ask a lot of questions about the context to the implementation partner, you can basically get to a point where there's like a meaningful wisdom of the crowds type of effect to guide that decision of “Is this context similar enough, is the sort of risk-reward similar enough where, you know, we feel comfortable as an organization providing advice to scale in domains that have not been explicitly tested?”. And I would say it's important to tread very cautiously and conservatively do in that type of, what would you call it? Almost like horizontal scaling rather than verticals? I'm not sure if that analogy makes any sense.

APPELT: No, I like that idea of thinking of the different dimensions of scaling. Yeah. And I would definitely agree that a lot of it is judgment calls. There's no mathematical formula for what's the limit of this result.

TREGEBOV: Yeah.

APPELT: Are there other evidence-based recommendations that BIT tends to make beyond just scale or don't scale? When you're closing out a project, what's the set of recommendations you tend to make?

TREGEBOV: Well, I think beyond some of the other considerations that I've already raised, one thing that I think is important that we can borrow a little bit from academic researchers, you know, whose methods and insights are sort of at the foundation of the applied work that we do is what's the next research question to answer? Is there continued iteration experimentation, trialing whatever the case may be, are the effect sizes that we've identified, do we think we sort of saturated the potential or do we think we found something that works better but not necessarily something that works best? I'm not sure that that fits the category of evidence based. I mean, I think it can be sort of evidence informed. I think that's the only other thing that comes to mind.

APPELT: Yeah. I like that because a lot of times at the end of our project, we kind of treat it as okay, wrapping that one up. But really it can be just the start of an ongoing stream of projects where you think about ways to improve or other things to test or combinations to add, kind of revisiting those discussions you have during the design phase where you were thinking about, you know, you had to pick a certain set of options, but maybe now you can open the door for other options.

TREGEBOV: Yeah, exactly. And, you know, this is like, it's an incredibly fun and privileged position that my colleagues and I have as sort of consultants and advisors and sort of, you know, independent enablers of this type of research. One of the frustrations is that our engagement sort of end. And I think, you know, especially if you're doing this work embedded within an organization, rather than as a sort of advisor to an organization, there's an opportunity to think about this work, not just in the sort of framework of a project, but as an ongoing program of research and experimentation. Like there's always a next, well, virtually always like a next question, and a next idea, and the next implication, to continue working on. And I think that's why, you know, we love the work that we do on sort of capacity building with organizations as much as we love the sort of discrete project stuff that we do.

APPELT: Yeah, absolutely. Same is true for us. And I think it just reminds me a little bit of a lot of times when we're either in projects or we've been coaching students on projects when we're in that design phase, we do say things like, you know, “This is just the first study. You're just trying something out here. You can always do future studies”. But then when we get to the end of the project, we're often, I don't know, whether it's just a sense of closure, but we just kind of, “Oh, put a bow on it”. And it is, yeah, super worthwhile to go back to that. I love that you've brought that forward.

TREGEBOV: Yeah. And I realize I may be taking down a sort of rabbit hole here, but there's something particularly important in that when you're doing classic sort of like nudge style interventions, either like as the sort of canonical Thaler Sunstein definition of nudge, or the sort of broader, you know, reference as like a low-cost intervention, which is that sometimes the barriers are deeper, right? Sometimes what you learn is that you craft some really thoughtful interventions and they don't get there and you're learning that the barriers are more systematic or that the interventions need to be better resourced.

There's actually a very logical and natural progression where, you know, I think a little bit about the work of Elizabeth Lenos, whom I had the privilege of collaborating with when she was our head of research at the Behavioural Insights Team in North America. And, you know, she works a lot on questions of sort of wellbeing and working habits and productivity of civil servants. And, she has this view now as an academic researcher at Berkeley of, you know, “Let's think about when we get this specific problem, let's start with is there a sort of low-cost solution that we believe, you know, scalable, low cost, easy to implement, solution that we credibly think might solve the problem?”

And let's start there. But let's have a whole plan, right? Let's have a whole program where, “Okay, we do that, works a little bit. Maybe it doesn't work. What's the next, what's the sort of next level of that intervention that's maybe a little bit more research, resource intensive.” That's such a compelling sequenced approach to me and sort of reflects that, what we were just sort of talking about, around like what comes next or what does scaling look like.

APPELT: Yeah. And I think kind of going back to what you said around capacity building. The other thing is that once you've done one project, a lot of times the data processes, the randomization, you know, that's all possible, the organization, the client has a bit more capacity on their side, so future projects end up being a bit of economy of scale of sorts.

TREGEBOV: Yeah, great point, totally.

APPELT: Awesome. Well, the other thing we often think about as we're closing an initial project and potentially considering future projects is how we're going to communicate out about a project. So how does BIT approach reporting? What channels do you typically used to communicate out?

TREGEBOV: I would say our reporting and communication strategies vary significantly. I would say, like, there are some aspects of our work that are very sort of consistent and repeatable, the trial protocols, you know, procedures that we use and others are very project specific, and I would say communicating and reporting falls into that second category. You know, we do like to have a somewhat consistent approach of a very sort of concise summary, a concise and consistent summary of the trial results, as much for our own internal knowledge management as anything else. At this point, we've just done hundreds and hundreds of trials as an organization. And so, having a record that you can sort of be comfortable almost like flipping through is valuable.

But in terms of the, you know, communicating more broadly, we like to blog, we've started doing more YouTube videos, we you know, we write two-page reports for our partners and we write 180-page reports for our partners. And I think it depends a lot on, again, in a perfect world, you mentioned this earlier, and I was sort of smiling because, you know, I had that down in my notes of like, again, you can approach these behavioural problems. What is the purpose of your communication in reporting? Hopefully it's more than just having a paper trail… some accountability mechanism or something.

And hopefully you are using that to produce a change, whether it's to promote scaling, whether it's to have other practitioners take up, you know, effective ideas in other organizations, other parts of the country or the world, there's probably a sort of purpose. And knowing what we know about human behaviour and the behaviours that we were hoping that those communications engender, let's be specific towards that. Are we trying to convince a very senior decision maker with very limited time and bandwidth to say, "Yes, well, let's have one, you know, killer graph or infographic or something?".

Whereas if, you know, we're trying to share something interesting that we found with the research community so that they can build on it, they're going to want to know everything. They're going to want to know, "Can I, you know, trust this result? Are the methods the same methods that I consider sound? Can I look at the sort of, you know, full regression tables to understand what the impact of certain covariates if my population is different?". Like it's just like they're such radically different use cases that I think, you know, logically there need to be radically different communications options.

APPELT: Yeah, and I think we're already getting towards my next question, which was around balancing different audiences because policymakers, lay audiences, practitioners, academics, like you said, have different purposes and uses for consuming. Do you find then that you have communications or reports or white papers or whatever it is that are intended for multiple audiences? Or is it usually each audience has its own kind of communication method?

TREGEBOV: Well, I think in a perfect world of unlimited resources, you've sort of developed products specific to an audience that's unfortunately not the world that we live in. I think, you know, sometimes you absolutely make compromises and you say, look, there, you know, we've got the time and resources to do one thing and there's four different audiences and we're going to try to signpost them. We're going to try to have an executive summary.

You know, one of the tricks, well not tricks, but like one of the things that I say, because this is something we struggle with a little bit as an organization because we hire really, really wonderful people who tend to have a really high level of intellectual curiosity and a deep love for research. And what they would find valuable and interesting in a report is not necessarily the same thing that our external audiences might find interesting and useful in a report. It's always like, "Okay, I know you want to share this. I agree, it's very important and interesting. Let's have an appendix".

You know, like I think like the appendix is one of the great things that we have. There are absolutely people out there who want to know, who want to track the, you know, 50 papers that you read and they want to go back and read those themselves. And you should make that learning available. You know, you should enable that type of deep engagement and learning. But if you're trying to achieve policy change, then, you know, often it's very important to have a clear, pithy, concise call to action, right? And then say, “If you want to know why we think you should do this, here are 60 lovely pages of explanation” or whatever.

APPELT: Yeah, totally. That resonates for every project I've done where you do a deep dive and you're just so jazzed about all the details. You want to share them out. Or when I'm teaching students and they just want to, you know, show everything they've learned. It is really hard to do that perspective taking. What does this audience really need to know? So, the appendix is our friend.

TREGEBOV: And it is really hard. And I hope what I said doesn't sound like pejorative at all, I think this is why you need to have different roles. I've been in the sort of role in my career where I've been the sort of lead author of the study, and you can't simply say to that person, “Well, what's the 60-thousand-foot view” to use a sort of awful consulting phrase or whatever, right? You know, business-y phrase.

But you need to have people that are working with you that are at least, you know, removed from the day in, day out of the project to say, “Hey, you know what, it really sounds like you're saying, is that we need to do more of these three things and less of those two things”. You know, one person can't, I don't think, sort of determine what all those audiences need, because the reason the work is so good is because people are fully invested and immersed.

APPELT: Yeah. And it reminds me of something I've always found, which is that it sounds like doing a short report will be easier because it's short, but it's actually doing the elevator pitch, doing that one pager that's really tough because you have to do that synthesis and that really selective picking of what information and you have to be able to do it all concisely and clearly. And so, in many ways, writing a one pager is harder than writing a 50-pager because you have to do so much of that curation.

TREGEBOV: Well, absolutely. I mean, if there's the same underlying body of sort of evidence, I would resource way more sort of time for a 10-page report than a 40-page report, because it's always going to start as a 40, 50-page report and you're going to need to go through multiple iterations. You know, literally we will assign more of people's time to a shorter report than we would a longer report.

APPELT: So, if you are forced because of time and resources and other things to choose kind of one method of reporting, is that a white paper? Is it some other kind of report? Is it a blog? What your choice if you're limited to one?

TREGEBOV: I would say like a very accessible blog or article with some good visuals, you know, for the sort of median project audience or however you would define that, I mean, it's a bit of a tough question, right? Because the answer that I want to give you is a non answer of 'it depends'.

APPELT: Yeah. It's tough. That's a tough one. Okay so I think you've already given us a lot of great recommendations around good reporting. Are there any things that didn't come up yet that you wanted to bring forward?

TREGEBOV: I think having somebody with distance, I'm just repeating myself, but having somebody with distance from the project, engage with your findings and provide advice on what'll be interesting to a given audience is very valuable. The same way that I feel like it's critical that your research protocols and theory of change are reviewed by somebody with expertise, but also with distance, I feel like that's quite important.

APPELT: Yeah, it's almost a way of pilot testing the communication as well as the design and everything.

TREGEBOV: Exactly.

APPELT: Do you have any favorite examples of good reporting, things that have struck you about what they've done in a way that has been compelling or useful or different?

TREGEBOV: I think in general, you know, I'd like to see the community move a little bit more towards the sort of pithy in effective communications. I have some colleagues that work in our Singapore office that have been experimenting a lot with the use of video as a way of explaining project findings. And I know I found that compelling. I'm not sure if that's publicly available, but it would be great to see a little bit more, because we know people are disinclined to read long documents.

It's like, you know, the first part of the EAST framework is easy. And yet we don't, you know, necessarily always consistently make it easy. And there are reasons for that, right? Obviously, the rigor and the depth are critical. And you do want to communicate that. It's not an easy challenge, but it's one that I feel like there's room for growth on.

APPELT: Yeah, I think that's definitely something that the field is working on and getting better at, but we're not there yet. Well, on the flip side of that coin, if you're thinking about, we've all come across examples of reports where we thought, "Well, that wasn't helpful". What do you think, what are the things that make a report poor? Like, what are common mistakes or things that we tend to not do well?

TREGEBOV: Yeah, so I think communications that are not aimed at a research audience that go to great lengths to communicate, you know, sort of methods behind findings. It's just it's throwing up a barrier, right? Like you're creating a barrier to your audience actually taking away what you're hoping to give them. Like, these reports are basic because we get to learn from them. Right, we get to learn things that are practical and relevant and interesting and most importantly, evidence based or informed.

And when you add a whole bunch of detail around the sort of methodologies that produced those findings or caveats, but not caveats that are going to be relevant, right, caveats that are factually true, but not actually relevant, for example, to the scaling question. You just you reduce that learning, right, from a really good place of wanting to be transparent, from a good and important place of wanting to be transparent. I've just seen sort of countless examples of that.

APPELT: Yeah, that is something I have seen as well. And something you mentioned there was what's relevant for scaling. And that was going to be one of my final questions on the reporting piece is, when we think about scaling, often the reports, the publications, the blog posts don't actually touch on whether the project was scaled so often as third parties or as a field, we don't have a good sense of what's been scaled or not. So how do you think we can better include scaling in and our knowledge exchange?

TREGEBOV: Yeah, it's interesting because I think the problem is that we do want to be timely as a community in sharing our research findings. You run a trial, you think the results of that trial may be relevant to diverse audiences, you want to get that knowledge to them. But the question of was it scaled or not? They come six months a year, three years later, depending on the scale and complexity of the scale of scaling anyways. So, I don't know. I mean, like, do you integrate scaling plans or considerations? I would say "Absolutely, yes". You know, put forward a plan for scaling.

But it's even more interesting to know, was it scaled? If so, how did it go? Both in terms of process and outcomes. And if it didn't scale, why not? What were the barriers that interfered? And so, I wonder, and again, this is just a little bit challenging for us, specifically because we are advisors. We are external advisers to implementation partners rather than the implementing organization ourself. It would be great to sort of issue, and we would need to find a better name for this, addenda, to say three years later, this is what happened. But I think that would be fascinating. I would be an extremely eager consumer of that sort of thing.

APPELT: Absolutely. It's been one of the interesting things of putting this module together where I've been trying to look at examples where, you know, things did and didn't scale and it's just so rarely covered in the reporting that it's hard to find examples. You can get like little hints, maybe. But it's hard to see this scaled, this didn't scale.

TREGEBOV: Yeah, but I mean, it would be very cool to sort of write into your research plans or agreements. And by the way, in two years, we're going to talk again, we're going to go through a sort of structured, you know, interview to understand what happened, what didn't happen, what the implications were. I'm sure there are researchers and organizations that have done this, I'm positive I'm far from the first person to you know, think about this stuff.

APPELT: Yeah, like you said, doing an addendum or an update, some kind of thing where it's just it gets attached to the initial report that, you know, three years later, this is what we know.

TREGEBOV: Because I think the academic community is actually starting to do this bit in the context of replications, which is not the same thing as scaling. Like, I think there's been this wonderful movement to, you know, redo studies and, you know, if sample sizes were limited or whatever to address those limitations in the original research, I think that that's been hugely valuable for the applied community, in particular, both for the academic and applied communities. But it doesn't really address the question that you raised. Did it actually even happen in these applied settings?

APPELT: Yeah, yeah. I think that's one of the areas where we're all still really learning. It would be great to do more knowledge exchange around that so we can all learn from each other as we're learning on scaling. Well, we're just about out of time, so any last messages or advice for our BI practitioners in training?

TREGEBOV: You know, again, I am so excited for this program. And I think, you know, it offers an enormous value not just to the learners, but to the sort of social impact and private sector communities more broadly. I'm happy to be able to participate in this small way.

APPELT: Well, thank you very much for joining. It was really, as always, a huge pleasure to bat around ideas with you, so I hope our listeners will find it valuable as they start entering the end stages of their capstone projects in the next few months. Thanks for joining us, Sasha.

TREGEBOV: You bet, thank you.

APPELT: And thanks to our listeners for joining another episode of Calling DIBS.