

One Knight in Product - E176 - Andres Glusman

📅 Thu, Jun 01, 2023 11:35PM ⌚ 37:46

SUMMARY KEYWORDS

experiments, tests, people, work, run, companies, good, learn, split testing, experimentation, traffic, signal, split, wrigley field, insights, andres, product, users, colour, number

SPEAKERS

Andres Glusman, Jason Knight



Jason Knight 00:00

Hello and welcome to the show. I'm your host, Jason Knight. And on each episode of this podcast, I'll be having inspiring conversations with passionate product people. If that sounds like the sort of thing you can get behind, why not come and join me and some of the finest product thought leaders and practitioners in the world on <https://www.oneknightinproduct.com>, where you can check out the back catalogue, sign up to the newsletter, subscribe on your favourite podcast app or follow the podcast on social media and guarantee you never miss another episode again. On tonight's episode we talk about doing the splits... no, no, not like a gymnasts split split testing that making sure that we get the best information out of our tests. What makes a good split test? Can everybody do them? Is it really possible to make an extra gazillion dollars just by making the Buy Now button a different shade of blue? For answers to all these questions are much much more, please join us on One Knight in Product.



Jason Knight 01:01

So, my guest tonight is Andres Glusman. Andres is a seasoned growth strategist and a pioneer in experimentation and split testing, turned company founder and CEO. Andres started out as a ballpark vendor at the iconic Wrigley Field where he had a front-row seat to the wisdom of crowds and presumably tried out loads of different outfits and signpost colours to sell as many 2 foot hotdogs as possible. But, he couldn't stay behind the counter forever and moved onto a career notably including nearly 15 years at Meetup, where he helped grow the company and LAUNCH it into the atmosphere and eventual \$200M acquisition by WeWork, and they all lived happily ever after. Andres is now the CEO of DoWhatWorks, a groundbreaking new platform that has analysed over 15,000 split tests and can hopefully FINALLY let us know what colour to make our submit buttons. Hi Andres, how are you tonight?



Andres Glusman 01:45

I'm doing great. What a wonderful intro.



Jason Knight 01:47

Thank you very much. I do have to ask though, was it hot dogs that you were selling at Wrigley Field or misrepresented your endeavours?



Andres Glusman 01:55

Yes, at Wrigley Field, my item of choice were the lightweight items so you could fly around because I was actually lugging all that stuff on my back.



Jason Knight 02:02

So hang on. then Was this unauthorised vending?



Andres Glusman 02:06

Oh, no, no, I was an official vendor as part of the Vendors' Union of Chicago and it was amazing. But no, I it was the kind of vendor that you walk around the ballpark with the items yelling and throwing stuff. And so when you sold Coca Cola or when you sold hotdogs, you actually have to carry around like a five tonne oven to like actually lug and deliver those hotdogs. So you don't want to do that. So I picked peanuts, ice cream, whatever is lightweight and easy to throw.



Jason Knight 02:32

Well, it sounds like it was character builder anyway, so that's the most important thing. But down to business, you are the Chief Executive Officer at DoWhatWorks. And we've talked about just now the 15,000 split tests and counting. But specifically, what problem are you solving for your customers with that analysis? And for that matter, what types of customers you solving it for?



Andres Glusman 02:51

Absolutely. So DoWhatWorks is a company that helps growth and product leaders do what works, right? So the problem that we're solving is a problem that I used to feel, which is essentially that when it comes to driving growth, if you can improve the conversion rates, on your website or on your ads, there is profoundly great things that can happen. But it turns out that improving conversion rate is super hard to do. When people try and do it, they basically are wrong, at least 80% of the time, even when they're running experiments. So the best people in the world run experiments to try and figure out what actually works and what doesn't. And even those people who have all this learning are generally not moving the needle, four times out of five. And what do what works is solving is essentially giving them the ability to learn from everyone else so that they can tip the odds in their favour.



Jason Knight 03:41

Yeah, I mean, those odds don't sound great. And we're gonna come back to those odds in a minute. But just in case anyone hasn't done this before, and you've kind of touched on it on a high level, but just in case someone hasn't done them or caused them something else entirely. How are you specifically defining a split test? Like what makes a split test for you?

A

Andres Glusman 03:58

Yeah, so there's lots of different descriptions of split test. Some people call them A/B tests. Some people call them split tests, some people call a multivariate test. Essentially, what it refers to systematically varying a user experience for different groups of users, and then comparing the end result to see how did the experience results in an a specific kind of activity? Did you see more people signing up when you had certain activity A when you when you presented certain language in the headlines? Or did more people sign up when you had B, which is why it's called a B testing? Right? Although we certainly see people are running tests, you know, as many as 789 way split tests when they really want to get crazy.



Jason Knight 04:35

Yeah, I mean, would you say there's a maximum that they should go for? Does that really just depend on lots of factors and not just the number.

A

Andres Glusman 04:42

The number one thing that it depends on is how much traffic volume you get. So the problem with running lots and lots and lots of variants, which is really fun to do very tempting, because in theory, you're going to learn lots and lots of things is that it takes a long time to learn. So when you are running split tests, it can take depending on the size of the company, But we work with six of the top streaming brands, for example. And even those companies that are really pretty darn big, are constrained, it takes them about a month to get enough traffic to run a specific experiment and get a statistically significant result. So if you're running three variants, four variants, eight variants, it can take eight times as long to get the result. And that's not necessarily what you want. So you really have to balance speed, with precision in the quest to actually figure out what to learn. Not just the fact that you can learn everything you need to figure it out, be very strategic about what you want to learn and how.



Jason Knight 05:31

Well, fair enough. And I can certainly think of a few CEOs and company founders out there that would in no way want to wait that long. So I guess, you know, short and sweet, right, but what made you start or decide to start a company that does this anyway? I mean, you work to meet up for quite a long time. Yeah, I think nearly 15 years through the we work acquisition, and I guess, maybe further divestment. I wasn't sure if the dates matched up. But ultimately, you

were there for a while. I'm sure you were doing a lot of growth there. I mean, I spoke to our mutual friend Giff Constable on the podcast a couple of years. Yeah, I love Giff. He's a great guy.

A

Andres Glusman 06:03

He's a wonderful, wonderful friend of mine. Yeah, I love Giff.



Jason Knight 06:06

Yeah, no, Giff's brilliant. I think he just lives up the road from you as well. So hopefully, you can go and have a coffee and say hello for me. But ultimately, he spoke about when I interviewed him, he spoke a lot about the, for example, when we worked over and there was obviously a much higher emphasis on growth and much more aggressive targets, or certainly that's what I remember from that conversation. So it feels like there was a lot of growth work going on there maybe before that, but certainly after that. So is that really where you kind of got the bug for all this stuff? Or did you kind of get this experimentation and script testing bug from somewhere else.

A

Andres Glusman 06:38

The bug, it turns out, I caught at the late 90s, when I first got involved with the online commercial internet in the late 1990s in the form, guys running ads on different portals back in the day, which I can get into if you want to kind of get your get your, you know, allow me to prove how much grey hair I actually have. But the bug for running experiments online and on websites really was something I caught at meetup. So you know, I helped launch Meetup and made the first \$14 of revenue, I had a variety different roles, including product and growth over the years. And we became early pioneers in the lean startup movement, because we just started experimenting with experiments early on. And we were extraordinarily lucky that the very first few experiments that we ran generated positive results. And that motivated us to want to keep going. And nothing is more addictive than winning. And so when we when you when you want to do more of it, and I ran lots and lots of experiments, our company started growing much, much faster, the more experiments I ran. And so I really fell in love with kind of the potential of what happens when you can run a lot of experiments, you certainly can grow a lot faster spending the exact same amount of money or less. Yep, the problem that I learned is that there's a lot of problems that occur with experimentation that nobody talks about, and you don't see people getting up on stage and celebrating the eight times out of 10 that they did not move the needle, you get them you get them on stage bragging about the few times that it did work. And so the problem that I said about solving was really my own personal frustration, which is when I was immune up, I doubled the size of my team and double the number of experiments I could run that grew to double it again, we certainly could run more experiments. But at a certain point, you sort of max out the number of tests you can run. And that maxing it out was not driven necessarily by my team size is really driven by how much traffic you have. And I also was sort of shocked to discover, like Man, there's so many more things I want to learn than I have the time to test. And so many of the things that you are testing don't actually work. And as a result, that sort of the the bug that I caught now, during that time I'd hang out with my buddies actually, it's so funny because gift constable and I and a mutual friend of ours,

Kareem, we all were the co organisers of the lean startup meetup in New York City, because we were all figuring this out at same time, then these became three of my best friends. And we meet up with those two guys. And then we'd also go hang out with the product leaders at Shutterstock and Etsy and various other companies in New York area that we're all really growing at same time, we would swap notes. And we were able to learn so much from each other and figuring out what worked and didn't work. And we'd literally come home from those sessions and the next day, start executing and be like, Okay, this is what the Shutterstock guys did. And this is what the people over at Etsy did. Let's, let's sort of borrow these things. And so for me part of the motivation, when I started, do what works was sort of saying like, Well, how do I recreate that feeling? How do I create the ability to learn from everyone at scale, without everyone needing to be in the same room without people needing to contribute their thoughts around it? And so it very much was on my mind there because I was absolutely such great, you know, had such a great experience collaborating with people and wanted to recreate that. And then when we actually launched it, it really started as a technology. We had a cool idea for this technology. And we my co founder and I were both playing around with ideas. We had both left meet up a little while earlier. We both left pre divestiture so we meet up was acquired by WeWork. And that afforded me some time to go play with different ideas. And my co founder had also left me up at the same time, he was an engineering leader that I had worked with for over a decade. And we had a great relationship. And he and I would get together regularly and play with ideas and toys. And we built this as a toy. And then we showed it to friends. And they said, they'd be willing to pay for it. And we sent them a stripe link and said, prove it. So they did pay for it. And that gave us enough confidence to go ahead and build this thing. And so we built the technology. And then we built the user experience that allows them to see the experiments. And then we've just had the great fortune of working with all these amazing clients to refine what we've been working on with them and in collaboration with them over the last several years to now have experiment discovered over 15,000 experiment. I think we actually just crossed the 16,000 mark this week. Wow. Which is great. I'm very happy about that. My numbers. Indeed, you can sort of see our trajectory based on on how quickly we're growing there. And that then gives us the ability to see patterns around what actually works and what doesn't work in order to be able to kind of tip those odds in your favour.



Jason Knight 11:02

But just to clarify, are you detecting people's split tests that they haven't asked you to detect? So for example, you can point a bunch of web crawlers at a bunch of different people's websites, and then automatically categorise stuff based on observed behaviour, or are you categorising the tests that people that are working with you have asked you to categorise,



Andres Glusman 11:21

We are categorising experiments based on what we're seeing through the technology. So the test that we've seen had been from over 1600 different companies that we're looking at on a daily basis, we're looking at over a million data points on any given day. And we're analysing them using our patented technology in a way that allows us to understand when the experiments are being run, and then ultimately, what's winning and losing. And then we've made it you know, we've really worked super hard to make the user experience our customers able to then tap in and be like, Oh, I wonder, I'm thinking about doing an experiment with social proof. Do these kinds of social proof work? Or is this just something that everyone does?

Because everyone thinks they should know? What are the different approaches that people might take in my industry. And so they're able to search by specific kinds of companies and by specific kinds of experiments, to be able to see the kinds of things that actually do or don't move the needle.



Jason Knight 12:11

I do have to ask, though, that, presuming that some people aren't asking you to check their own tests, and you're kind of just calling them that some people kind of object to you tracking their split tests, or that they don't really care?



Andres Glusman 12:21

It's funny, because when we first started this thing, the we always got two questions. The first question was, Is this legal? Are you allowed to do this? And we, we were very careful early on to make sure that yes, it is we don't look good in orange. So we paid back when we were really, really just getting started, we had these tiny little customers, you know, just just really good signal. We paid a lot of money to lawyers to make us make sure that we were doing everything in a legit way that is completely aboveboard. The second question we get is, can I pay you to not be in your database? And we said, No, that is not something we want to do. We don't want to be in the shakedown business. So it didn't feel very good to us. So yeah, we want to be in the business of helping people learn from everyone. And so the reality is that even the biggest of companies that we're tracking, that happened to be also our customers, they learned much more from everyone else than they're actually contributing in to the system. Although, of course, you know, the only way that an experiment gets into the system is by our engine, detecting it, and working through its process. So there's nothing that we might know from a friend of ours or from anyone that influences what goes in, it's entirely driven by what the engine is seeing.



Jason Knight 13:28

Sounds exciting, maybe I'll sign up. But are there any top insights that you've come up with, like you just talked about analysing across all these different companies, or the 1000s and 1000s of tests that you're doing? I'm sure you're doing some analysis of your own as well, some sort of meta analysis so that you can get almost like a State of the Nation of people's tests across your entire database. So I guess the question is, are there any certain specific design or interaction patterns or things that people are doing on their websites, on a kind of aggregate level that reliably give good results? And maybe we should just stop fighting, trying to change how people do those things? And they should just do that thing every single time?



Andres Glusman 14:08

That's 100%, what we're all about. And that's kind of the core question. So what's interesting about our services that we launched it, and we were able to deliver insights to people, so visit data, and they're like, that's really, really cool. I'm amazed to be able to see this one experiment from a competitor. And then they said, Well, I want to see how does, how does that

pattern hold across multiple people like this? And so they started doing it themselves? And they said to us, well, hey, can you just give us recommendations and give us an analysis on our specific pages? And so the answers are very much dependent on the kind of company that we're working with the kind of sector that they're in and the kind of page that they are looking at. I mean, what's really fascinating though, is you can see certain trends emerge. And then you start looking at experiments that are related to that trend. So some people start actually testing it. So a year ago, there's a Superbowl ad, and the ad was by Coinbase, and had just a QR code floating around the screen. Very famous. And then everyone says, Oh, we got to start using QR codes and our advertising, right, we got to use QR codes to try and drive drive results in and get people to sign up. And so what we started seeing soon thereafter is a lot of experiments on heroes, on their signup pages, right trying to encourage people to use a QR code to download the app. In theory, that should really speed up the process of being able to get to the opposite. One of the classic problems that everyone who's ever had a mobile app is trying to bridge the gap between a computer and a phone. So we saw one company tested, we saw two companies as we saw three companies test and it's almost like universally true that every single time we saw somebody tested in this specific way, in the hero to the Logout user, it did not win it lost and lost and lost. And so those are the kinds of patterns you can see over and over and over again, on the kind of a new front, which is really cool, because so many trends and so many patterns and really conventional wisdoms that you see people embrace. They're embracing it for all the wrong reasons like that ad was a success, because there was nothing else like it, not because there's a QR code. And you see people do this over and over and over again, on all these different emerging trends. And the reality is most things that people take as a given are not really backed by data. They just start doing it because the guy or girl before them started doing it. And you see that over and over and over again.



Jason Knight 16:16

Well, I mean, losing is obviously very on brand for crypto these days. But I do think it's interesting, this whole idea that people, they go down the obvious path because of the cognitive biases that make that sound like a good idea. And the kind of well known immunity to statistics, or the inability to grasp or understand statistics that pretty much everyone has somewhere they kind of have to train themselves to get good at that. So QR codes aside, are there any other things like key insights, where we'd sit there and say, for example, always make your purchase button blue or something like that? Or is it a little bit more nuanced than that? Because that's the classic example. Right?



Andres Glusman 16:54

It's extremely classic. The very interesting thing is one is it's very nuanced in terms of what to do. What's fascinating, though, on the since you brought up button colour is, it's sort of the classic thing that you get your hands on a testing tool. And you're like, alright, what can we what can we change? Like, let's change button colour? Because that's actually the easiest thing to change. There couldn't be anything easier to change than button colour on these tools.



Jason Knight 17:20

I've certainly worked some developers that would argue if you're in my time, but you know, let's just assume those are outliers.

A

Andres Glusman 17:25

Yeah. And so if you're running a test you like, let's run it on button colour, what we see is that everyone tests red versus blue, green versus yellow, blah versus blah, right? Does, it doesn't really matter what the outcome that we see is that almost universally the case that that's not a factor that matters, the words matter. The size matters, the shape matters, the number of CTAs matter, the approaches that you're taking, that are the headline around it, the call to action, whether it's sentences, or bullet points, all those things matter. But in their nuance to the specific kind of company and this specific kind of challenge you're working on? For sure. But sadly, the number one most popular thing that everyone experiments with is a bit of a dud, which is actually super sad, because what could be more soul crushing than having your first experiment at the gate? Be this massive loser over and over and over again? And you're like, oh, did we just make a terrible investment when we invested in whatever technology we were using to make it easy to run test?



Jason Knight 18:20

Yeah, good luck getting authorization to make the next test as well. Exactly. But that's an interesting point, though, right? Because there are some companies out there that experiment quite a lot, you've talked about some of the stuff you were doing back in the day. And of course, there's gonna be a bunch of people that you're detecting and working with now that are also doing the same, but also a bunch of companies, probably quite a lot of companies that don't really do very much testing at all. And I wondered if you'd spoken to people out there when you're out and about talking about this stuff that maybe you're in that situation where they're not getting to do too much of the stuff. And if you're able to call out some of the main barriers or blockers to companies doing more experimentation to help drive some of the growth and the kind of like strategic areas to go in, that you talked about earlier.

A

Andres Glusman 19:05

When it comes to running experiments. There's some extremely good reasons to not run experiments. And there's some really bad reasons. The...



Jason Knight 19:15

This sounds like an A/B test!

A

Andres Glusman 19:16

It is a bit of an A/B test. And I've lived through all of them. And what often happens is that there's a pendulum. And so when you're very, very, very small, and you don't have very much traffic on any given page, running an A B test, will take you three months, five months, 10 months to get a result if you have a very limited amount of traffic. Yep. And so as much as you want to learn as awesome as it would be to learn that this thing to drive a 15% lift, do you

really want to do the work to wait 10 months to get the results? So traffic is one of the biggest drivers of getting of running experiments or not. Similarly, there's other constraints that people feel like in regulated industries. You got to get your lawyer involves you gotta get your regulators you got, there's so many people that have to say yes to getting a test out the door. And so there's a lot of constraints that slow you down. That being said, in those industries, small improvements can make a pretty big difference in the number of people signing up for your credit card, for example, that's kind of a huge deal, the number of people signing up for your direct to consumer health product is a pretty big deal. So it just more constraints. And so that slows it down. But the impact there is pretty profound. So they have headwinds, but they're able to do them and but it's just slower and harder. The case that you often hear, which I think is sort of the wrong approach, is I don't want to run an expense, when you have the right amount of traffic is I don't want to run an experiment, because it's going to be it's going to slow us down, we're going to know if it worked or didn't work after we launch because we're gonna just see a big lift, or we're not. And if it's not a big lift, then we're gonna, well, no, but what ends up happening is that the thing gets launched maybe a little bit faster, because it's not run as a split test. And then immediately, everyone hovers around a computer, usually around an analyst computer and say, did it work analysis, I have no idea if it worked or not. And so this port analyst now goes, runs around for months feeling executives breathing down their neck, trying to figure out whether or not this pre post analysis actually made a difference in what way and try and unpack it. And it's so hard to decipher, that ultimately, this gain that you thought you're getting in terms of being a lot something faster, you sort of measure the right thing, which is the time it took you to know if it worked. So from idea to actually knowing if it worked. If you measure that time, it's probably the same amount of time or slower, but your precision and your ability to learn, just got just fell through the floor. So the argument that says, Oh, well, we can launch a lot faster, we don't have time for this stuff, when you have the traffic is really not a very good argument at all. Now, I can tell you like, like an argue against myself, I can tell you all the times when I've learned the you should, in fact, not be running experiments. And there is such a thing as running too many experiments. And I've been there too. So you know, when we were really running, start and run more and more experiments and meet up and we were seeing when after winning when it's really feeling good. Suddenly, we're like, oh, let's learn everything. Let's not test anything out the door. And we need to make sure that every experiment would make my biology teacher proud from high school. And so we basically then, you know, only change one small thing. And then another small thing. And then another small thing. And that was one thing. So did we learn with a lot of precision? Yes. Was it worth the trade off of like, operating on smaller and smaller increments? No. And so at a certain point, what happens when people fall in love with experimentation is that they start running too much. And there's a sort of a Goldilocks is sort of the the parable there, right? You want sort of you want to be running at the right aperture, you want to run on things that are going to make a difference you want to run so you actually know you made a difference. You want to assemble building blocks together, or things that work so that you can move the needle and make bigger swings. But you don't want to overdo it.



Jason Knight 23:02

Yeah, this reminds me of some of the discussions that you have with people about product discovery in general, like quality of discovery as well, where people get into this kind of we've got to speak to the next user, we've got to speak to the next user. So it's just like that on an industrial scale. I guess it's just, I guess it's just this interesting tendency. And I've seen this and spoken to other people about this before. And this idea that product teams are where good

ideas go to die, because they have to spend so much time kind of going in these loops, to learn everything that they can that I forget to actually ever ship anything. So I can imagine that does frustrate certain types of people.

A

Andres Glusman 23:36

That's right. And what you're really looking at there is people wanting to not be wrong. More than you want people wanting to make a difference. I mean, obviously, everyone's hearts in the right place. And everyone who's working on something generally finds himself in product, because it's really, really cool. You get to affect behaviour. Yeah. When people start getting scared of being wrong, when people start being scared of like, oh, it's not, it's not super scientific to the point where you know, it can pass the 99.9% confidence level. Well, at that point, you're being way too cautious. In you're letting science replace judgement. Like that's, that's such a bad way of saying it, you're letting you're letting a desire, you're letting a desire for experimentation, replace judgement, and experimentations inform your judgments, they make you smarter, they make you better at your job, they make you more able to hone in on what works. But even the most precise experiments even, let's just go with a 95% confidence interval. You know, by definition, there's a margin of error in a 95% Confidence Interval test, and you have to live with error and errors are part of life.



Jason Knight 24:39

Yeah, well, don't tell some of the people I've worked with but let's talk a bit more about that risk. Then and you said before this call and actually join us called as well that 80 to 90%. I think a split test don't beat control. If they don't perform any better than the original version that you were trying to improve. And you said then you would never take those off. is a casino, which I agree with, I would not play that game. But how can we stack the deck in our favour and attempt to run more successful tests in general?

A

Andres Glusman 25:11

Yeah, those are impossible odds that you would never take at a casino.



Jason Knight 25:15

Not even in Vegas!

A

Andres Glusman 25:16

Not even in Vegas. You know, those are the odds that venture capitalists take and why did they take them? Because they sort of banking on the winners to more than make up for the losers want, right? Yep. And choose, you just have to be very smart about screening, what you're going to do or not do in order to maximise the odds that you get that two out of 10, the three at a time the five out of 10 win, right? Because those impacts can be very, very large when you actually start stacking them up. And they do compound on a singular flow. So it's hard, but

worthwhile is actually the number one thing to note. The second thing to note is just to say, Okay, well, how do you get your hands on signal? Before you figure out what you're going to test? Gotta get hands on signal, when you're generating ideas? How do you get hands on signal when you're prioritising? How do you get your hands on signal when you're thinking through all the different assumptions on how you're executing it and the different ways in which all the language comes together and the layout and the design and the elements and when you include a How It Works section or not, the more you can get signal, the better. Now, of course, I should say that everyone should use do what works, get that signal net be awesome. And that would be the best signal, the best signals always come from your work. So it'd be awesome if I could say that actually probably do believe that. But it's not the only place that you can get signal. You can get signal from teaming up with other product leaders like I used to do you can get signal by running those qualitative experiments and just watching people use stuff, getting qualitative feedback as a way of triangulating in Oh, that's interesting. I had a conversation with I showed them, this new thing we're thinking about, I showed them a rough mock up of this thing. We're thinking about it, had them try and use it in this weird way. And it was a signal to me, and they actually not only were willing to pay for it, they gave me their credit card number. Okay, that's a pretty good signal that there's a demand for this thing, right? There's ways of getting signal all across the board, and you just have to be clever about understanding whether or not there's more reason to do it than to not do it. And if there is, well, then you're tipping the odds in your favour. And that's kind of the good news about terrible odds is that they're not that hard to beat. If you if you can, if you can, if you can really figure out a way of getting that incremental. It's almost like Moneyball, famously from the Michael Lewis book



Jason Knight 27:18

He gets on base! Back to the Wrigley Field, right?



Andres Glusman 27:20

Exactly. He gets on base. It's just funny, because I grew up playing soccer, I didn't play baseball at all. But I do love that movie. And I love the what it's all about, which is sort of around using a very simple heuristic, a simple way and a few ideas with data upfront to be able to sort of outdo your competitors, because you've got the information advantage. And so that's really what it's kind of all about is how do you get your hands on data before running the experiment, in order to have the edge on everyone else.



Jason Knight 27:47

Don't run too many at the same time. But I was speaking to a friend recently, in fact, today about split testing. And she was wondering out loud whether they could ever really scale up properly. She said that she's run a bunch of A B tests in her career. And even if she considers once on the same product, that if she'd have kind of taken a winner from all of them should be about 40,000% up already because of the different kinds of signals that she's got back from those tests, as she also a pined that maybe there's more of a limiting factor of saying, Well, we're using these tests to make sure that we don't make bad decisions, rather than that we're making like knockout new good decisions. Do you think that that's a risk of running too many

tests that you end up kind of optimising for just not making things worse? Or do you think there's always a potential golden egg out there somewhere that you can win, and as a way to operationalize these tests in a way that can honour those into costs,

A

Andres Glusman 28:39

Doing no harm is a good starting point. So that's not a bad spot. I don't think that's why we're all in this business, though. I don't think that's why people get into product or get into driving growth or make products and or run companies, the goal is to try and affect people and to try and change your behaviour for the better. The experiment ended of itself? Well, first of all, it's not a bad thing like like getting a 40,000, or whatever, she said, 4,000% 40,000% lift, I think a lot of people would sign up for that in a heartbeat. So that's not bad at all, just in terms of conversion rates. The essence of it, though, is what do you learn about humans, every single experiment gives you an insight. And, you know, the joke around experience is experience is what you get when you don't get what you want. And so all those losing experiments that you see people run that we're doing harm, or they're avoiding doing harm, or at least learning, but the ones where you're actually doing right, it gives you the most profound signal about what people are responding to, what's the attention there, what what is it that ways in which people behave? And so if you can be strategic about harvesting your insights, so that you're not just getting a numeric win, but you're getting a transferable insight, then you've just like, using baseball analogies, and you've just hit a homerun, right, you've just you've just like, Grand Slam if you want to go all the way there, right? So you, you basically have just compounded the benefits that you're Receiving across the board. And it's not that the methodology is good or bad, it's the mishap, it's the appropriate application, because when it's applied properly, then you're overcoming the natural human shortfall. That is to believe that your assumptions are correct. And you're figuring out which ones are right and wrong faster. Which ultimately, is the difference between like kind of an okay, career, and just a gangbusters career, right things feeling like they're like coming out, we have wind at your back or wind at your front. And that's really the big difference is how well you are able to learn, and how quickly you're able to incorporate the lessons into the next thing you do.



Jason Knight 30:37

I think there's always an argument that you can always do one more thing to check it out, as long as you're not, as we talked about earlier, doing continuing to get into a loop of that. Like, there's always something new to learn. But I guess you've just got to work out when the tipping point is, I mean, it's maybe going back to the casino, rather than baseball was for a second, like when to hit and when to stick, right? That's exactly right. So like, When can you go in confidence,

A

Andres Glusman 31:02

When there are diminishing returns at a certain point you need to move on, right, you need to think about a different component, there's only so much juice, you're going to squeeze out of out of the lemon air. That's That sounds so negative is only so much juice, you're gonna squeeze out of whatever fruit you got going on there. And so you got to just sort of know when there's diminishing returns, and then move on to the next thing.



Jason Knight 31:19

But speaking about juice-less lemons, I mean, we talked a bit about the fact that you need to have a certain base size that you can go out to certain number of users that you can go and experiment on or experiment with. When I think of split testing, I'm thinking of big mass market, B2C apps with billions of users. And yeah, we talked already about the Google 50 Shades of Blue and all the money they apparently made out of that. That all sounds great if maybe slightly inaccurate, for every use case, but I'm a B2B guy. And I've worked on some pretty nice products in my time with a handful of users in the hundreds rather than the 1000s, or millions or billions. They all paid a lot for the privilege, but there certainly wasn't enough people there to start doing what didn't feel like there was enough people there to do any kind of test on. So is there any hope for us b2b types? And will we ever get to run a split test?



Andres Glusman 32:11

It's interesting that you said some of our most active customers are actually in B2B, not B2C. And the reason for that is because there is now consumer like expectations with business level traffic. So they have the real real challenge of delivering a beautiful experience that feels like it's as polished as using Netflix, but they have just a fraction of the traffic there. And that even less traffic on a super important page, like a pricing page, even fewer people get there. And so what are you gonna do, the opportunity to learn is still there, and we see it. And we see a lot of experiments in b2b. So it's actually one of the areas that we we really love looking at, it's really fun to look at. And the challenges are kind of the the need for data upfront becomes that much more profound. But we do see a lot of people experimenting with it, we do see patterns, and we do see the ability to kind of optimise and learn. I don't think anyone can. I think everyone if they had the choice would rather be able to leverage insights and go forward with what works or doesn't work. And if they had the traffic, and it wasn't hard to launch the experiments and want to do those experiments themselves. And so just the need to be judicious with that traffic, you have fewer chips in front of you. Casino going right, you have a certain that sounds



Jason Knight 33:33

Like me at a casino!



Andres Glusman 33:35

Yeah, you have your chips in front of you. And so you have to be even more judicious on when, when you got to hot deck.



Jason Knight 33:40

Yeah, that makes a lot of sense. I guess also one of the things kind of going back to that experimenting with people as well as try to do what you can to, you know whether or not this is exactly the same, but maybe recruit people to take part in kind of a hybrid of experimentation and more research based stuff. So you can maybe get people to do some task based stuff with

you're kind of getting them to come in and try to do things within your app and just try and use whatever data going back to that squeezing the lemon thing again. Just try and get whatever you can out of them in some way. Like not everything has to be a statistically significant, whatever confidence interval based on hundreds of millions of users, right, there's still something you can do. Most of the time, I guess is the takeout.

A

Andres Glusman 34:19

I could not agree with you more. It is profoundly important in that scenario, and it's important to remember that you're just triangulating your way into a truth. And there's no one magic bullet surveys are great, but they have a limitation. usability studies great they have a limitation split testing great. They have a limitation guess what everything has a limitation that you put enough of them together there, they start to blur each other out and the kind of the peaks make up for each other's valleys and you can basically start to hone in on a truth if you need to be 99%. Right? I really hope you're putting a medicine in somebody's body or you're trying to put a human being because if you're trying to like have that standard and have no error whatsoever at that level, you really, there better be some really significant consequences. Because the flip side of it is the speed with which you move is going to be glacial, which is why it takes however many years it takes to get a medicine approved. And that's a good thing. That's not a bad thing.



Jason Knight 35:21

Don't get the anti-vax people started!

A

Andres Glusman 35:23

Exactly. I should be careful with my analogies here. But it does take a while to get, you know, to get to get everything designed to put a human being up in space, that's a good thing. It's a very high consequence. If you have billions of people visiting a certain page and you want to change everything for everyone, well, that's a pretty big significant change, it will have a huge impact. So maybe you can run all the small scale experiments to gather data beforehand, and you want to be very precise about it. That makes a lot of sense. You have 15 people visiting your website, or you have a smaller number of people coming through, and you want to try and drive as much results as you can in this short period of time as possible. You're gonna have to give on that precision. And in return, you're gonna get the ability to like, run more things and try more things, which is going to be again, that flipside of win rate times the number of things you try, is your outcome at the end of the day. And that's what you need to optimise for.



Jason Knight 36:10

Well, there you go. There's hope for us B2B people. Well, plenty of food for thought there. And hopefully something that could inspire people to maybe at least think about doing some split testing, maybe even look at your platform and check them out as well. But after this, if they do

want to reach out to you to chat about do what works, or find out more about split testing in general, or maybe see if they can tap you up for any Chicago Cubs memorabilia. Where can I find you?



Andres Glusman 36:35

So my website is <https://www.dowhatworks.io>. The place where I tend to hang out online is LinkedIn, I find it's kind of the the kindest, and the most productive and these days positive environment these days. I really, I enjoy the positivity there. But you can find me at you know, just do Google search for a LinkedIn search for Andres Glusman.



Jason Knight 36:56

Well, I'll make sure to link all in the show notes anyway, and hopefully you'll get a statistically significant number of people heading in your direction. Wonderful. Well, that's been a fantastic chat. So obviously really glad you could spend some time talking about some really interesting topics around experimentation and testing. Hopefully, we can stay in touch. But yeah, that's for now. Thanks for taking the time.



Andres Glusman 37:15

It was my pleasure. Great, great conversation.



Jason Knight 37:19

As always, thanks for listening. I hope you found the episode inspiring and insightful. If you did again, I can only encourage you to hop over to <https://www.oneknightinproduct.com>, check out some of my other fantastic guests sign up to the misters Skype on your favourite podcast app and make sure you share your friends so you and they can never miss another episode again. I'll be back soon with another inspiring guest but as for now, thanks and good night.