The Measure Pod: #35 Is tracking too much a bad thing?

[00:00:00] **Dara:** Hello, and thanks for joining us in The Measure Pod, a podcast for people in the analytics world. I'm Dara, I'm MD at Measurelab, and I'm joined as always by Dan who's an analytics consultant also at Measurelab. Hey Dan, how are you doing?

[00:00:29] **Daniel:** Yeah, good thank you. Yeah, how have you been?

[00:00:32] **Dara:** Yeah, pretty good thanks. Looking forward to diving into our topic this week, which is, is tracking too much a bad thing?

[00:00:40] **Daniel:** Yes, no, I don't know.

[00:00:42] **Dara:** Finished, short episode. So this has come up a few times, hasn't it? Well, a few times on the podcast, it comes up a lot in our day to day work, but it's come up in a least two of the previous episodes where we've basically talked about the idea of people potentially tracking too much, and then not being able to make head nor tail of the data they're collecting or potentially not trusting it because it's been added to over a years and years as different people have got their hands on the keys to analytics. So we figured it would be a topic in its own right that we should dig into a little bit more.

[00:01:13] **Daniel:** Yeah, absolutely and it's been brought front and center when we are working with people now on a GA4 migration plan. So quite often, when you talk about a GA4 migration, it sounds relatively simple. You apply the tag, you do the events, you kind of transition the reports. There's way more to it, and of course it's a different product and there's some learning curves there to get through. But actually one of the biggest issues or the bits that take the longest time are actually migrating the events, the tracking events, all of the customisations, the things that you've done for your business alone, those non-standard in Universal Analytics.

[00:01:45] **Daniel:** What we quite often find is that within a legacy setup, this may be added onto over years and years worth of, of it being in existence. You get to a point where there's such a big setup, there's so many different events, there's so many things tracked. So they are tracking too much, objectively too much. And when it comes to migrating to GA4, it becomes quite anxiety-inducing because it feels like such an insurmountable task of how the hell are we going to take these 100 events and these 200 custom dimensions and metrics for the 360 customers. How are we going to turn that into something that's completely different in GA4?

[00:02:19] **Daniel:** So it's just more of a conversation, not objectively is tracking too much a bad thing. I suppose it depends will be the answer in all 99.9% of the cases. It's more, as we're now approaching this with a fresh set of eyes thinking, you know, we have to move over to this new tool. We've got a fresh start, we've got a clean slate. And actually there's a really good opportunity to say, were we tracking too much before, and do we need this going into a fresh state ecosystem as it were.

[00:02:43] **Dara:** Well, exactly that. And it's kind of forcing our hands, forcing everyone's hands in a way to look at the current implementation and think, is this fit for purpose? Do we really need all of this event data that we're collecting? And it might be in some cases that obviously you'll need additional tracking. So there might be interactions on the site or in the app you're not tracking currently that you do want to track, but there's almost certainly going to be a bunch of stuff that's been in there from something in the past, an old version of the website or old functionality.

[00:03:11] **Dara:** What we've often seen as well is people will track something for a specific reason. Maybe if you've got some kind of interactive content that's on the site temporarily, and then you forget to take the event tracking off afterwards or. Sometimes people will use a tool like Monetate or something like that. And it'll fire events into GA (Google Analytics) and you don't necessarily use the data in GA (Google Analytics) itself, but it continues to track indefinitely unless somebody goes in and says, what is this? Who's using this? Let's get rid of it. So the timeline with GA4 is giving everybody a reason to do a bit of a spring clean and look at the implementation and think is this data that we're collecting, is it all being used? Is it all being used in a meaningful way? And if it isn't, then it potentially doesn't need to get migrated over to GA4 at all.

[00:03:52] **Daniel:** You mentioned one of my biggest pet peeves with anything within this industry, this career, this job, and that is implementing something and walking away and never revisiting it. So much priority is given to when we need this new tag put live, when we need this event tracker put live when the new website goes live, or we're launching a campaign next week get these Floodlights set up in GTM. But when it comes to then removing and once the campaign's done or removing them once the features no longer there or revisiting it at a future point, or even thinking about giving people access to these platforms, everyone's quick to add people in and as a trainer, people add me into their accounts all the time, so I can train on their data. But then I say to them today, remove me from your access. Remove me tomorrow, remove me whenever and I very rarely do, if ever.

[00:04:33] **Daniel:** So there's always that case of people are quick to add in, but not think about whether they should remove, it's just not part of that process in most workflows. And it, again, it kind of basically builds up and up and up to the breaking point or to a point where you have to migrate it and now all of a sudden you're scratching your head and thinking, what the hell is this for? Who implemented this? Why is it relevant? Just to kind of give a bit of advice, I suppose, around how I approach these things. And I always ask two questions and I suppose we can apply it to the GA4 migration situation, but also it's just more general.

[00:05:01] **Daniel:** First of all, I ask, is it needed to be tracked? So is this tracking needed? And if the answer is yes, I ask a second question, which is, is it needed in GA (Google Analytics)? Because they're two very different questions, is the data needed? Yes. Well, can you get it from somewhere else? Is it already somewhere else? Having it in GA4 in this case is a completely different question because I think I've mentioned on this podcast a number of times, but my go-to example and one that I have quite a lot of conversations about is things like refunds in Google Analytics, you can integrate refunds into GA (Google Analytics) so that you can measure your purchases, how much had been refunded and obviously accounts for the negative revenue to kind of balance out the totals, that kind of thing.

[00:05:39] **Daniel:** But the reality is, you've got that. You've got that data. You must have that data, right? Otherwise it'd be a terrible job of dispatching and delivering all of your items and refunding them. So there is a stock management system, there is a backend system that has that data, the factual, real-life data. GA (Google Analytics) does a very bad job of tracking anything exactly, right. There's always a, there's always a variance, there's always a difference in the data set. You can then start to try and clean it up, but it's almost like an impossible task. You'll never quite get it to be perfect. And then I think about, well, what is the point of GA (Google Analytics)? GA (Google Analytics) is a marketing tool. It's purpose is to measure marketing effort, digital marketing efforts at that, understand if they convert on the website, do a bit of attribution and figure out how much value that is generated for the ad that you just served, probably a Google Ad let's be honest.

[00:06:21] **Daniel:** So when I'm thinking about Google Analytics and the amount of effort and time and energy it takes to build a feed, to feed in things like refunds into GA (Google Analytics). Is that going to change how you're advertising? Is that going to change what you use GA (Google Analytics) for? Because it's not going to be the correct numbers in terms of stock and revenue, and if you're not going to be optimising your ads or, or doing any analysis around the attribution from it, then it's almost like a vanity project. So there's two questions, is it needed? And then is it needed in GA (Google Analytics)? Two very valuable questions to really figure out, do I need to take this event from Universal Analytics and put it into GA4.

[00:06:55] **Dara:** Refunds is obviously a classic example, but there's other things as well that you see people do and I'll get myself in trouble here so I'm going to add a disclaimer, I'm not saying don't do this in GA (Google Analytics) because there are use cases for it but I've often seen where people go overboard with tracking menus or tracking scroll depth on every page. And again, I'm not saying don't do it, but there are other tools that may be provide the type of insight you're looking for in a better way. So if you're looking for kind of qualitative information around how people are navigating through the site, there's potentially better tools for doing that. Are you really going to look at that event-based click-through data on your main menu, or are you really going to look at scroll depth in detail? If the answer is yes, then fine do it at GA (Google Analytics), but often what people actually want to know is what content are people looking at on the page or focusing on, or if there's two links on the same page to the same destination URL, which one's more likely to get clicked.

[00:07:51] **Dara:** These are the things you can answer with Hotjar or any other heat mapping based qualitative tool. And you don't necessarily need to have all that data in GA (Google Analytics). So loads of examples of this, and we fall into this trap as well. When you're used to using GA (Google Analytics) day in, day out, you think, well, I could kind of shoehorn this data into, and then I've got everything all in one place, but the advice really is use the right tool for the right job and I guess that's the key point. And almost every account we look at it has suffered from this at least at one point in time where somebody has thought, you know, let's shoehorn this data into GA (Google Analytics) because we can, and just because you can do it doesn't mean that you should.

[00:08:30] **Dara:** The other thing I was thinking about there as well is, this doesn't just apply to GA (Google Analytics) tags is not just about event tracking, but you could add lots and lots of tags into your GTM container over time or hard coded things onto the website that might've been useful at one point in time. And you might not even be paying for those services anymore, but you might still have the JavaScript on your website, or you might still have a tag in GTM, and I know some of the GTM audits that we've done over the years sometimes there's hundreds of tags that get removed at the end of the audit because the website owner is thinking, I don't know what half of these are, or we used to use them once upon a time, but we don't anymore. So this applies way, way beyond just GA (Google Analytics) but it's looking at what you're tracking on the website across the board and, you know, is there anything on there that's adding bloat to the page that's potential complicating your use of cookie banners or gaining consent, if you're not using it, why make life difficult for yourself? Why have it on there in the first place. Go through and do a sweep and get rid of anything that you're not actively using.

[00:09:32] **Daniel:** That's a really good point. It definitely extends beyond just Google Analytics, but I suppose where playing devil's advocate, putting myself in a website, owner or client shoes. It's also then, well, I've got my day job. How am I going to remember to go into GTM and remove tags and things like that. And there are ways of doing it, but actually it becomes really difficult if your job isn't to own the governance of this product. If analytics is always a side, a side aspect to your job, that makes it very difficult to do and unfortunately it puts me in this hypothetical situation at a greater risk, because especially as you mentioned with things like governance around cookie consent, or just consent in general to comply with, things like GDPR around where we are, or other requirements globally.

[00:10:12] **Daniel:** We don't have the luxury of not knowing or not caring or not investing time and energy into it anymore. We can't just put a tag live, that's not how tracking works nowadays unfortunately. So although the tool exists for you to be able to click and deploy a floodlight or a GA (Google Analytics) tag or a Google Ads remarketing list tag, it doesn't mean you should right. Yeah, of course you could, but should you, and it's then it goes into this kind of bottomless pit of like, well, how do I know how to do that? How would I set this up? How do I configure all this stuff? It becomes very difficult. Then we're basically saying, by the way, remember that you did this thing six months ago and take it off. Obviously we can say, oh, the tools allow you to add notes and you know, annotations, and you can use your calendar, but it's like the reality is quite often it might be outsourced. And if it's outsourced, then we're kind of relying on a third party to remember to do this.

[00:10:56] **Daniel:** For example, if you've got a marketing agency that is setting up all of their own tags in your GTM container, I don't want to speak on every agency's behalf, but quite often, they're not going to be GTM experts. They'll be marketing experts or marketing tag experts, and it's not in their interest to not serve the tags to everyone, right to measure the performance of that campaign. So things like cookie consent and consent in general is going to be not high on their priority list. So unfortunately, it always comes back down to the client, it always comes back down to the business, whoever that representative is to be aware of that. And I think even if we're saying anything right now, I know we've drifted slightly off topic from event migration, but it feels like it comes back to just being aware that these things exist, and I suppose just governance in general.

[00:11:35] **Daniel:** When we talk about data, is tracking too much a bad thing? Yes it can be, if you're tracking too much in the sense of it's just overwhelming or unnecessary, but also are you tracking too much that you really shouldn't be, you know, are you getting data without consent or without legitimate interest? And these are all different aspects of is too much data a bad thing, and in some cases 100% it's a terrible idea. And in some cases it's just an inconvenience, so there's a spectrum there.

[00:11:58] **Dara:** Yeah, and I think it makes me think of the kind of trust issue. It's such a common scenario for a client to come to us and say, we don't trust our data because we've got all this stuff that we're tracking that we don't understand. We don't know who added it, it was added years ago and sometimes you almost do want to just kind of clear it all away. And back to the kind of the original point, we're talking about there being a good opportunity with the GA4 migration, it's a chance to do that. It's a chance to think, hang on, what are we holding onto here that we don't really, we don't really need. And you made a point there that it's maybe not somebody's day job to control that and to go through and wade through all the layers of analytics.

[00:12:36] **Dara:** Another reason why people might be reluctant to let things go is fear. Fear that they're going to lose something that maybe somebody is using. And I'm sure we've said on the podcast before that if it's not showing up in any of the main dashboards that are used or any main reporting that you know about, try switching it off and pretty quickly somebody is going to come and say, hey, where's that metric that I need, but, you know, chances are, if you're not aware of where it's being used is there's a good chance it's not being used. But I think that that's probably a driver for all of us where we think it's fear of letting it go and then realising you needed it. And just thinking, I'll just hold onto it just in case.

[00:13:12] **Daniel:** This is digital hoarding right? It's hoarding, but in a data context, I might need it just in case or it's started to happen now we've seen it where people are asking, oh, I need to export every single piece of data from Universal Analytics just in case. And I'm like, well, in the last 10 years you haven't needed it, so in the next 10 years you probably won't either. Especially as it becomes less relevant, but it's like, well, I have that fear that I might need it and I won't have it. So actually thinking about it, and I may have say something controversial, Dara, and it's just that there is nothing business critical in Google Analytics. There is no business critical data in Google Analytics, but the thing is, these are all nice to haves and from people that work in GA (Google Analytics) day in and out like ourselves, it's a bitter pill to swallow, but there's nothing that the business cannot do if they didn't have Google Analytics.

[00:13:57] **Daniel:** I mean, case and point, you're talking about consent management and cookie management. If people don't opt in, you can't track them and it hasn't broken. They haven't gone under, businesses haven't gone under because they can't get consent to store an analytics cookie. And obviously there are hundreds of other alternatives to Google Analytics that aside, but the point is, is that in a sense, there is nothing business critical in Google Analytics, or at least they shouldn't be. So if you do have something five years down the line that you think, oh, I wish I kept that, it's not going to make or break the business or the decision at that point in time.

[00:14:25] **Daniel:** I think we spoke about before, actually in one of the previous episodes, but you'll always be able to access that historical data, the aggregate data, the totals, and exporting that into a database or a spreadsheet even is going to be fine to do, but you don't need every single event and page view and hit from Universal Analytics to make sense of something in the future.

[00:14:43] **Dara:** I think another point as well as this isn't, it's not a binary decision and the question isn't, should you track a lot or should you track a little. The question is, can you track too much, or can too much tracking be a bad thing? And the answer is yes, it can be. Especially in the cases that we're kind of talking about, where people don't have a full understanding of the data they are collecting, or it's collecting data that was useful at a point in time, but isn't useful anymore. So if you don't trust the data that you're reporting on, because you don't have a historical understanding of all the different layers of tracking that have been added by different people, then that's a bad thing, and definitely time for an audit and definitely time to clean up your implementation and make sure it's tracking the data that you need and you're going to use, and that you can understand it and everybody in the business understands it.

[00:15:28] **Dara:** It's less about tracking a lot or a little, because you could track a lot of data and then you could control who has access to what so you could through dashboards or through other forms of reporting. You could give access to different people within the business or different third party agencies. You could give them access to the data that they need and only the data they need. And if you, as the website owner, or as the analyst, if you want to have every interaction on the website tracked, then that's fine. The issues come when it's just a bit of a free for all and nobody really knows which data points they can trust and which ones they can't trust because there's no full understanding.

[00:16:04] **Dara:** So the question isn't is tracking too many things a bad thing, it's more about is the way that you're tracking and your understanding of it, if that's based on lots of different people adding different things at different points of time, then that's probably going to lead to distrust and that's going to be a bad thing. It's not to say that tracking too much is inherently bad because it's what we do. We track, and if it's useful or could potentially be useful because you've obviously got the issue with GA (Google Analytics) of it's a bit late to realise a week after your campaign went live, that you should have tracked it. So, you know, it is sometimes worth tracking something in advance, as long as it's well-documented and everybody understands what it's doing and how it's working then track as much as you want really, within the limitations.

[00:16:51] **Daniel:** That's it and I think it all comes back down to the idea of governance and control and documentation, which again, the luxury of time to be able to do that is not available to everyone. So I suppose my advice at least would be, if you don't have the luxury of being able to invest the time and the governance of this product or these products then play it safer on the safer side and don't over track to cause this overwhelm or confusion or issues, potential issues down the line. If you have someone that's a full-time employee, you know, an analyst that maintains Google Analytics, then in a sense, go to town as long as you can absorb the information and you know what you're doing and it's all well-documented then why the hell not. But yeah, governance is a key aspect of all this stuff.

[00:17:33] **Daniel:** Slight left turned Dara, but I want to play a bit of devil's advocate, so Universal Analytics, let's assume the free version of Google Analytics here. There is a hit limit at of 10 million hits per month don't need to tell you, but a hit is anything that is traditionally a page view or an event. There are other things, but quite often we just think of those two things. So let's say I'm a big website and I'm getting, I don't know, 5 million page views a month and a couple of events thrown in there per page or whatnot, and I'm kind of edging towards the 10 million. Then you get to GA4 and you think, oh great there's no hit limit. I can finally do the tracking I want to do, why would I not just track everything? Why would I not just track every event, every interaction and because there's no limits within the tool and there's no sampling within the reporting interface, the reporting workspace that is. Why would I not just go to town and track every button click, every interaction and then figure it out afterwards?

[00:18:21] **Dara:** I mean, I would say again, back to what I mentioned before, if you're likely to use it, then do it. Don't do it if it's tracking for the sake of it, and it's going to lead to confusion and if different people are setting up different events and there's no shared understanding of what you're trying to do, or there's no shared taxonomy for how you're setting these up, but if you're going to use it, then by all means, that's what the tool is there for. And again, you can control how much of that data you share with different people through dashboards or the kind of reporting. You can control it from somebody just receiving a metric every week, right the way through to you as the analyst having full access to all of that data in BigQuery.

[00:18:56] **Daniel:** What's the opposite of devil's advocate? Devil's antagonist? Either way I jumped back over the fence, but the other thing I would warn myself if I was in that position or asking those questions would be around. Yeah, there's no monthly hit volume limits within GA4, but there are other plenty of limits, you know, there's still an enterprise version of this product and they still limit it in different ways. So be mindful that although this one limit has gone, or this one limit has changed, there are other limits and new limits introduced actually that weren't obviously present in Universal Analytics. With more events that you track, there's going to be more event parameters and more user properties, more data, more contexts that you want to pull into the platform. And so things like custom dimensions and metrics are going to fill up pretty quickly. And when we do come to use the Explore workspace to do those custom explorations and those custom reports, there is sampling, there is 100% sampling. So the more events you do send in the, you are going to be limited in terms of how much value you can get out of the product. Yes, there's built in standard reports, there's going to be no sampling for sure, and you can use the API and pull that data out, whatever, but actually querying that data for that kind of deeper analyses, you're still going to be shot in the foot, you're still going to be at a disadvantage.

[00:20:05] **Daniel:** 100% is your prerogative. If you're set up in a good position to use the data, to know what you're doing and to document it well, go to town and yes, you will have that cap of 10 million hits a month removed, but know that if you all of a sudden decide to go crazy and it's now 100 million hits a month. It's actually going to make the use of the product very hard because of these other limits you'll come up against and the sampling that's available in the Explore workspace. So do a bit of a research, have a play within the demo properties before you jump headfirst and decide to go crazy.

[00:20:35] **Dara:** We're getting towards probably summarising here. I think we could probably pull it back and re-stress the point about using GA (Google Analytics) for what it's intended for and this kind of covers quite a few of our points around, you know, avoiding the risk of shoehorning data in, or trying to track too much in GA (Google Analytics) or over rely on it. When effectively it is a marketing analytics tool at the end of the day.

[00:20:58] **Daniel:** Yes, exactly. It's way more of a marketing tool than Universal Analytics actually was, or at least from my opinion. Universal Analytics was from my perspective, a data product, data tool that became and evolved into a marketing product, you know, through the acquisition of Urchin from Google through over the next 10 years of Universal Analytics and it felt very distinct. It was a data product and you could layer on these advertising aspects into the Google ecosystem. GA4 has been built from the ground up as a marketing product within this ecosystem. There's inherent things that are automatically shared like audiences and conversions within the connected ad products. So actually reminding ourselves that this is a marketing tool and its primary focus is marketing attribution basically and kind of measurement of the success of those marketing touch points. So kind of going back to my two questions, is this data needed? If so, is it needed in GA (Google Analytics)? Is it needed from a marketing perspective? That's basically what we're asking. And so if it is then great let's put it in, otherwise there might be other alternatives, there's other tools out there that might do a better job or a more specific job of trying to get what you need.

[00:22:01] **Daniel:** If it's a marketing focused thing that's really important to you and for your campaigns that you're running, then sure, pull it into GA (Google Analytics). I think safely, you can probably park and maybe not migrate a bunch of stuff from Universal Analytics. I think in most cases, most migrations that I've worked on or been part of, we've actually not migrated a lot of stuff and it's been absolutely fine. Be deliberate with it but actually not migrating stuff is a really good opportunity to kind of do that clean slate we've been talking about.

[00:22:30] **Dara:** Yeah and to put my own spin on those two questions. I think if you intend to use the data or somebody in the business intends to use the data and if GA (Google Analytics) and this is a big if, if GA (Google Analytics) is the right tool to collect that data, then there is no such thing as too much tracking. If on the other hand, you trying to use GA (Google Analytics) to do things it shouldn't do, or you're trying to track a whole lot of extra data that nobody's going to use, or nobody's going to understand then yes, of course, that can be a bad thing. It can lead to confusion and distrust, which is obviously the worst thing for an analytics implementation because if you can't trust the data, then nobody's going to use it.

[00:23:06] **Dara:** All right on to our favourite bit of the show where we fake having really interesting lives. What have you been doing outside of work?

[00:23:14] **Daniel:** The bit of the show that enables people to skip ahead or to turn it off. What have I been doing? And so we just, as of the time of recording, we just had the double bank holiday Friday and Monday off, which has been, you know, a long weekend, which has been awesome. There's a friend of mine that has been building an indoor skate park down in the town that I live in. So although I didn't really help in the construction of the ramps and things like that, I managed to help out, sort out the outside and sweep up and move some furniture and all the kind of mundane, small things that don't involve heavy machinery or cutting wood. That's my forte, I said I'm a soft handed computer worker or something like that I think I got called when I was there. I got lumped with the kind of clean-up duty, the point I'm trying to make is I had a really good time, had a little skate at this indoor skate park and yeah, I contributed even if it is in the smallest possible sense. How about you Dara?

[00:24:02] **Dara:** Well, I've been putting my soft computer hands to some actual, probably a bit of an exaggeration to call it manual work, but I've been sorting out my garden. I took advantage of the nice weather and so I spent as much of it outside as I could just tidying up. But yeah, not exactly hard labour, but yeah, it is exactly, it was harder than typing on a computer that's for sure. But yeah, outside enjoying the long sunny weekend, just making the most of it.

[00:24:27] **Daniel:** Awesome.

[00:24:28] **Dara:** Okay just before we wrap up, where can people find more about you Dan, or if they want to get in touch how can they do so?

[00:24:34] **Daniel:** Sure, [LinkedIn](https://www.linkedin.com/in/danielperryreed/) is the way forward. So just such my name, links in the show notes of course, but also my website [dananalytics.co.uk](https://danalytics.co.uk/?utm_medium=podcast&utm_source=the_measure_pod&utm_campaign=35). How about you?

[00:24:43] **Dara:** You'll never get tired of saying Danalytics will you?

[00:24:46] **Daniel:** I really don't and I think I came across another name analytics. What was it? I can't remember, but I think there's almost like, I feel like I have to create a group of people, I think we have to form a little posse. I think that'd be quite fun, or maybe the opposite.

[00:25:02] **Dara:** I don't yet have daralytics the website, but you can also find me on [LinkedIn](https://www.linkedin.com/in/darafitzgerald/). That's it from us for this week, as always you can find out more about us on our website, you can find all previous episodes of The Measure Pod over at [measurelab.co.uk/podcast](https://www.measurelab.co.uk/podcast/?utm_medium=podcast&utm_source=transcript&utm_campaign=35). If you missed any, or if you want to go back and listen to your favourite episode or episodes again. Otherwise if you want to suggest a topic or even if you want to come on The Measure Pod and discuss that topic with Dan and myself, you can reach out to either or both of us on LinkedIn, or you can email us podcast@measurelab.co.uk. Our theme music is from [Confidential](https://open.spotify.com/artist/4erl5USs7RuiUU4QeiEQFO?si=xyGxD84PQHO78TKlvmZKLw), and we've got links to their Spotify and Instagram in the show notes. Until next time, I've been Dara joined by Dan. So it's a bye from me.

[00:25:48] **Daniel:** And bye from me.

[00:25:49] **Dara:** See you next time.