Welcome to the second video in this research series. Today, I'm going to be talking about key ingredients of a reproducible research plan. This all centres around the question of, on the topic of thinking before you do, as Laura mentioned in the introduction. So the aim of today's lecture is to introduce reproducible publishing methods and to outline a reproducible research checklist that we've created. And we hope that the content discussed in this video and the next one will help you approach your own research study with a reproducible mindset and also offer some practical advice to start adopting these reproducible research methods. At the very heart of all of this work is the idea of transparency. And one way that we can be transparent is to use a checklist to guide how we approach our research and design our research outputs. So what does reproducible research like look like? Well, there are many different forms of research and it can vary depending on the research topic. But broadly speaking, reproducible research tends to fall into two categories, especially in the publishing context. So the first is a pre-registration, and this is a timestamped read only version of your research plan that you created before you begin your data collection or data analysis. And this is normally uploaded to an online platform such as the Open Science Framework or another publicly available repository. This can be made available at the very start of your research project, or it can be put under an embargo so that it becomes available later, maybe once your results have been published. Then the second main format is a registered report. And in many ways this is very similar to a pre-registration. However, peer review happens in two stages in a registered report. So in many ways you create a very similar data analysis plan, but this undergoes peer-review journal and before you undertake your research. So as you can see here on the slide, and the peer review process is split into two stages. The first being after the study design, and then the second is after you have run your analysis and are undertaken your data collection and then undergoes peer review for a second stage. But I will discuss these in more detail or a registered reports in more detail in a separate video. So today I'm going to focus on preregistration as I feel like the topics within this can be applied to both the pre-registration and a registered report. And to introduce this topic, I'd like to use the metaphor of a recipe. When we approach cooking or baking, we will use a recipe often and in many ways it's important that we know exactly the quantity of each ingredient and also the order in which these steps are our completion because it's all well and good knowing that you need eggs and butter and sugar in a cake. But if you don't know the order to put them or to carry out these steps. And then my cake could look very different to your cake is applying a similar method when we approach research is really important because a lot of the time when we look at research that's published in the literature, there isn't enough detail to reproduce the research recipe or the data analysis plan. We hope that this reproducible research checklist will serve as a guide to help you make your research more transparent and more reproducible. So the first kind of starting point in this, in any research journey is thinking about, okay, what's my general research question? What is the gap in the literature that I've identified and how your, why do I want to research this topic and how am I going to do so? Relatedly, this then informs the hypotheses of your research study. It's very important that these are specific, concise, and testable hypotheses. Following on from that, you have your research question, you now have your hypotheses. But one of my variables of interest. So what, how am I going to measure the construct of interest in my research study? And how will they be measured? And what are the different measures available? Do I have covariates? So you may have your predictor and your outcome variable, but are there other factors that might influence that relationship That's our important to take into account in this research study. And it's very important at this point that there's a clear rationale for including each of these covariates. What's my sample

size? Whereas my sample size, whereas my sample going to come from and what are the features of this sample that might influence the inferences I can make about my results and the outcomes of this research project. Then importantly, what are the statistical tests that I'm going to use to appropriately test these hypotheses. And these may seem like quite straightforward questions. But often when we read a research paper there isn't enough detail or there's a lot of detail. There can be lacking for each of these points, which again, would make it very difficult for someone to come along and follow this recipe and try to reproduce the findings. Also, Dacia is, has, has many different features beyond the primary conflict of interests. So how are you going to treat missing data? What happened? What does the data look like? What's the distribution? Are there outliers? These are all important things to think about when you start your research project and also thinking about these in advance before you've looked at your data can again help avoid things like p-hacking and hacking. Laura mentioned in the introduction. Then also importantly, what criteria will I use to make inferences about the findings of my data? Often, as Laura mentioned, the researchers tend to really focus upon having a significant p-value of less than 0.05. But what about the effect size is, what does that tell us about the relationship that we're trying to study? And what about the confidence intervals? These are all aspects of your research that you should be thinking about in advance of undertaking any analyses. And also, it may feel quite daunting to try and get your head around all of these different concepts. But especially for maybe junior researchers. Or if you haven't done a statistical analysis of this kind before, really creating space within your research study to try and understand these topics can really make you a better researcher so that you have a thorough understanding of why you're undertaking research in this way and how you're going to do it. And this will help you later on in the research journey. So again, returning to this question of, does my recipe have enough information for someone else to make the same research study? And this doesn't necessarily mean that the findings have to be the same if they're not stars in itself is interesting. But again, we should have the tools to be able to try and reproduce the research and other research studies in the field. So why should I care about pre-registration? And it sounds like an awful, awful lot of work. But yes, you are front-loading the work, but your future self will. Thank you because as I already mentioned, you taking the time to understand these research topics, these statistical methods will really give you a solid foundation in understanding why you are undertaking research in this way and the different decisions that you encounter in the lifecycle of the research project. And it'll be very helpful for future work. So e.g. if you have a pre-registration that can form the introduction and the methods section of your paper. And also especially for a Ph.D. students are masters students. And a lot of the time, this can be writing that goes directly into your thesis. So e.g. I. Undertook a registered report, my PhD, and a pre-registration, and I was able then to use all of this, all of this writing in my thesis. So it really does save you a lot of time down the line. And also when you get stuck into your analysis, it is difficult too. Stay, stay focused on, on the task at hand sometimes. And you can often become confused about, okay, well, what am I actually doing here? We're having a pre-registration means that there's always a blueprint that you can refer to and this can be very helpful for your future self. I promised that your future self will thank you. So there are many examples of pre-registration available online. The Open Science Framework is particularly good for templates and examples, and these can be applied, these pre-registration can be applied to data collection based studies, secondary data analysis, and also qualitative work. Methods and structure of these may vary, but the principles are the same. So often the data that you include in your pre-registration will be different depending on whether you're collecting your own data or analysing a pre-existing data set. But the idea of being

transparent and including as much detail as possible remains the same. And in the next video we're going to look at an example of a pre-registration and kind of explored the checklist in a bit more detail. And hopefully this will give you a further understanding of what is involved and how you might use this as a template for your own work.