

Title: Getting Started with Open Science: A Panel for Trainees and Mentors
Date & Time: TUE, JUN 08, 2021 12:00 PM – 01:00 PM EDT

APA's Peggy Christidis: Hello, everyone. Welcome. Thank you for joining us today for this AP webinar on getting started with open science, a panel for trainees and mentors. This webinar is part of a new APA series called the essential science conversations, where panelists and audience members have the opportunity to have an open dialogue about various topics in psychological science.

APA is the world's largest association of psychologists and is strongly committed to a large number of activities designed to improve, promote and disseminate psychological science. For example, APA is extremely active in federal advocacy to protect and increase our scientist's ability to do our work. APA communicates psychological science to the general public. It promotes our science to a wide array of agencies and disciplines and APA creates opportunities for essential dialogue that moves our field forward. My name is Peggy Christidis. I'm a Science Program Officer in the Science Directorate of APA, and I will be serving as moderator for today's essential dialogue on open science.

Today, we have a great group of panelists, but before I introduce them and get us started, I first need to go over a few housekeeping details. First, the webinar's being recorded. Once the webinar ends, everyone who registered will receive an email with a link to the recording, and you should receive that email shortly after the webinar ends. I estimate about maybe three hours after.

We'll also have this webinar recording available on APA's YouTube channel. If you weren't aware, APA does have its own YouTube channel. Just go to the YouTube app and type American Psychological Association in the search box and it'll take you to this webinar as well as numerous APA webinar recordings that are available for viewing. Many of you submitted questions for today's webinar when you registered, thank you for doing so, and we'll try to get to as many of those questions as possible. You will also have the opportunity to ask a question as the webinar is taking place in real-time.

There's a panel labeled questions in your go-to webinar dashboard, towards the middle or bottom 1/2 of the dashboard. Just click on that panel and type your question into the box and we'll be monitoring those questions throughout the webinar. Without further ado, I'm pleased to introduce our panelists for today who will be discussing open science with us and answering all of your questions.

First, we have Dr. Katie Corker, who is an Associate Professor at Grand Valley State University in Allendale, Michigan. She's also the executive officer for the Society for the Improvement of Psychological Science or SIPS for short. Second, we have Dr. Nick Fox, who is a Research Scientist at the Center for Open Science, where he works on better understanding the markers of credibility in the published literature as part of the darkest score projects and we'll have him talk a little bit more about that later.

Last, but certainly not least, we have Ms. Ummul-Kiram Kathawalla who is a PhD candidate in counseling psychology at the University of Minnesota's Department of Psychology. She is also currently doing her internship at the VA Boston Healthcare

System. Let's have a conversation about open science. Nick, Katie, and Ummul welcome to the webinar. We're happy to have you here today and we can all turn our cameras on. I will also turn mine on.

Katie Corker, PhD, Grand Valley State University: It's ready.

Peggy: Okay. Hi, welcome. Good to see you all.

Ummul-Kiram Kathawalla, MA, University of Minnesota: Hello.

Peggy: Great. Nick, we're going to start with you. We can't talk about open science without talking a little bit about the Center for Open Science. We were lucky enough to have Dr. Brian Nosek here Co-Director of the Center for Open Science. He took part in a webinar that we did back in April. He was our first guest in this series. We'll talk a little bit about some of the things he discussed. I want you, Nick, to tell us what you do as a research scientist at the center and tell us what it does and specifically, a type of work you do there.

Nick Fox, PhD, Center for Open Science: Sure. Thanks so much for having me and all of us, super happy to talk to everyone. The Center for Open Science it was founded in 2013, like you said, by Brian Nosek from the University of Virginia, and the overarching goal of the center is to increase the openness, integrity, and reproducibility of scientific research. We do that in multiple different ways.

One part is by developing tools to help researchers increase the openness and increase the reproducibility of their work. One of them that I'm sure we'll talk about later is the open science framework, the OSF, we also support preregistration and we support preprinting of manuscripts. That's all on the product side of the center. I'm involved in the research side of the center where we try to understand the research of research.

Right now, like you mentioned, the score project, we're trying to better understand what about research signals its reproducibility and credibility. Can we actually find markers in papers and as part of a very big multi-lab project that's been going on for over two years now that includes computer scientists, that includes psychologists, that includes researchers, and other social behavioral fields, the replicates group you may have heard of as well. We're just trying to better understand what's there in the research

that helps signal how papers may be credible.

Peggy: Fantastic. Thank you. Ummul, I want to talk to you about a paper that you wrote called *easing into open science*, and that was recently published in collaborative psychology. We'll actually put the reference to that article in the chatbox in just a moment. In that paper, you provide a lot of guidance to graduate students and their advisors on how to start engaging in open science practices. Before we go into the details, I wanted to know what motivated you to even write this article?

Ummul: As Nick just said, thanks for having me, I'm glad to be a part of this panel. This paper really was something that emerged from the beginning of grad school. Early on in my grad school career, my advisor is Dr. [unintelligible 00:07:13] at the

University of Minnesota. We became really interested in learning more about open science and we began that journey together. I went to the conference of the Society of the Improvement of Psychological Sciences for a couple of years in a row and then joined the executive board as a graduate student member as well.

Through those experiences, I really submerged myself and started to integrate a lot of practices into my own work. A lot of the graduate students started reaching out to me and asking me, "Oh, Ummul how did you do that Preregistration? Or how did you do this? I realized that I was pointing them to different resources, writing a lot of emails, and started to be like, wait, this is getting too much. So then I created a little how-to document and started sharing that from my grad school friends. From that point, Moyne and I discussed great could actually be something that's useful to get out there. That was the birth of the paper.

Peggy: Well, I recently read it, I found it very helpful and I highly recommend it to everyone listening in today. I put the reference in the chatbox for everyone. Before we talk about ways to include open science practices into our professional lives, I think we need to talk about some of the challenges or perceived challenges of doing so. Katie, you wrote a chapter for a book called the *Portable Mentor*, where you talked about the need to have an open science workflow in order to have more credible and rigorous research.

In that chapter, you mentioned that much of the process of research is actually hidden away and that makes replication of research findings problematic. Can you talk a little bit more about this problem and why it exists?

Katie: Yes, absolutely. When people do research, the typical way that it was done prior to the last say, 10 years or so was you might get an idea for a study, talk about it with some local colleagues, design it, do the study, and then write it up for publication. What ultimately shows up in that final publication is really a summary or just an abstract or some people even say an advertisement for the research that you actually did.

All of the details of what actually went on in the lab, when the participants were there, what steps did you actually take to produce the results? What analysis strategies were used, all of these different details are things that are considered too much information for your typical research article. All of those details are really crucial when it comes to replicating other research, or when it comes to trying to even reproduce what's been done in a published article. A lot of the Open Science movement has been about opening up those details and providing people with opportunities to share those pieces of their work that might've been previously only known to the local collaborating team.

Peggy: That's great. Nick, can you elaborate more on this? When people contact the Center for Open Science, what are some common questions or concerns that they have regarding engaging in Open Science practices?

Nick: A lot of what I hear, both when people contact the center and just giving talks, having the flag as being someone from Center for Open Science get a lot of the same questions, which is, "Yes, this all sounds great, but what can I actually do?" I think that's the whole purpose of this webinar. There's so many different steps, and I

think a lot of it comes down to comfort and also internal usage. It can be great to talk about things in the abstract, where it's like, "We should share for reproducibility and we should share for the betterment of science."

It's also good to be a little bit greedy. What is it you get out of it? For me, I'm a very forgetful person and I lose a lot of things. For me, it's so much easier to put my data and my code, and my methods all in one place. That way, future me in two years, which I've done now, I'm editing a manuscript and this had to come up, where it's like, "Oh yes, what happened? How did I do this? How did you do that?" You can just type it in and there it is. Beautiful. I know where everything exists.

Another thing can be preprinting. I want to have as many eyes on my work as possible, and Katie mentioned this too. It's not just about a local collaboration at a university, or among a lab that my lab works with. I would love to get lots of different opinions, lots of different experiences looking at my work saying, "This is strong. This is weak. What do you mean when you say this?" I have a certain perspective and I come to my work a certain way and maybe that biases my writing a certain way, and I would love to get as much eyes on that as possible.

Preprinting a manuscript and getting it out there before it's published, helps in that editing process to say, "Oh, well, I've had this blind spot because of my own experiences or my own way of approaching this work. Thankfully, because other people saw it, not just the three or five people in peer review, not just my lab, I've student for years, I can approach this in a fresh way." Again, we can talk about all these different aspects, but I think thinking about what the internal motivation is for behaviors, and what you can get out of it, is one way to find that first step. Then from there, move forward into other behaviors that maybe other people find more encouraging for themselves.

Katie: Peggy, your question made me think about the pragmatics of the situation. I get this question a lot as well when I'm giving talks. Again, all of these ideals sound great in the abstract, but what do I actually do? Ummul's chapter and my chapter are both good resources for getting a sense of the different types of practices that are available, and then pragmatically, how do you actually get started?

I usually recommend people just pick one thing, one practice, and start with that. It can be really overwhelming when you hear about all of these different techniques, preprinting, sharing data, open code, pre-registration. There's so many different things you can do. Just pick one of them and start with that. Get expert in that practice and then you can slowly add additional practices as you become more comfortable.

Peggy: Yes, that's a great idea. Starting with one thing and then moving forward. It's always the hardest thing to just get started. Once you take that first step, I think things start to flow a little more easily. Ummul, I wanted to include you in the conversation as well. What do you think is one of the biggest challenges that graduate students face since you're currently a graduate student? You're almost done, but you're currently a graduate student. You're on the ground there. What do you think is the biggest challenge that you hear from other graduate students when they want to start engaging in Open Science?

Ummul: Starting with a more graduate student's specific concern is, if your advisor is not on the same page as you would want to integrate some of these practices and change the approach in the way that they approach different projects. In the article, we do talk about different ways of discussing some of these issues with an advisor. I understand that there's sometimes limitations for that.

There are different things that are in the paper and that we're even talking about so far today that I think that graduate students themselves can individually do to start moving in the direction of Open Science that are smaller little behavioral changes. Like working on a project flow for yourself, or working on how you write your reproducible code. Just as examples that are things that are more on the individual workflow.

Something else that I think is not only graduate students-specific that is a barrier that sometimes comes up is just feeling like you're going to do it wrong, or that the only way to do it is to do it all. Just like Katie had also mentioned, I really do think that Open Science doesn't mean that you have to do all of the different behaviors in order to be considered moving in direction.

Even doing just one or thinking about these ideas and noticing ways that your actions cannot be reproducible, or potentially might be leading to some of these problems that we-- That birth the movement of Open Science, is steps in that right direction. The fear of doing something wrong, I think is something that's a myth, so to speak, that we can think about in a new way of just taking baby steps, or taking a step.

Peggy: Great. I want to keep going with this topic. We covered a little bit about first steps, but let's dive a little more deeply into it. We got a lot of questions about this, and I'm going to read you a few questions from our audience members that they submitted. One person asked, how can a graduate student get involved in Open Science? Where is the best place to start?

Another person asked, where can I learn more about best practices for Open Science as I move toward becoming an independent investigator? The third person asked, for one who wants to engage with Open Science practices, but hasn't adopted them yet, what's the best thing to do first? Is there a simple first step? Katie, I wanted to start with you. If you were starting a project, what would be the first step you would do?

Katie: Oh, boy, that's a great question. Ummul's idea about starting with the practices that you can control things in your individual workflow, I think that's a great idea. Concretely, the first step that I would probably start with is the Open Science Framework, I think. Using the Open Science framework as a way to organize your materials can be a really nice way of thinking about what you want to share and when. One issue when it comes to open and transparent practices is, at what stage of the project do you want to share something?

One of the nice features about OSF is that it has really nice public and private features for each different project and component. That means that you can work privately in the way that you're used to, and then later make the components or the projects public so that others can see what you've been up to. That can be a way to

gradually work more openly. Eventually, you might feel comfortable just keeping everything public, to begin with, but as you're getting used to it, that can be a really nice way to figure out, which pieces of this do I want to share and at what stage?

Peggy: That's great. We'll talk more about the Open Science Framework in just a little bit. Nick, if you started a research project, what's one of the first things that you would do?

Nick: Now, given that I've been doing these behaviors for a little bit, I would probably start thinking a preregistration. A preregistration is basically just thinking about your research plan and writing it down before you start executing your research plan. Thinking about, what sample size do I need, and what am I basing that off of? Is there some power analysis? Where am I basing the observations out of?

What materials am I going to use? Am I going to use a survey that is published and out there, and do I need to get access to it a certain way? Things like that. Thinking out the plan that you're going to execute and then writing it all down. The great thing about preregistration is that there's not really a wrong way to do it, it's all about getting it out there first. At least, on the OSF, there are many different types of templates for preregistration. Some for new data collection.

There's going to be a new template coming out soon for existing data, or secondary data analysis. This is something that was also born out of a SIPS Conference, which is exciting. There's also an open-ended preregistration as well, so maybe you don't know all the details but you know some at this stage and putting that down. Again, a small step, a first bite to get you into the mindset. That's something I would start with.

Peggy: Great. We'll talk about pre-registration in more detail later on today as well because we've got a lot of questions about that. Ummul, you laid out in your paper specifically some first steps in getting started with Open Science and I like how you gave each step, a difficulty rating, ranging from easy to difficult. Can you tell us about some easy first steps?

Ummul: Of course, our ratings, not everyone might agree but this was-- They're a little arbitrary but based upon our experiences, I'm glad that that was helpful to think about the different behaviors. Some of the ones that we labeled as easy was one that the first is creating a journal club and learning more. You would ask some folks that asked about how to learn more about Open Science and you think like this webinar is a way to do that.

One of the things that I found helpful in grad school was creating a journal club on an area that I wanted to learn more about and having a community of people to read articles and discuss helps motivate me also to learn about something new and having people learn together. That was one of the behaviors that we put as easy and there's some resources in the article because you don't have to reinvent the wheel, there's a lot of folks that have put together reading lists and things to be able to learn. Using the resources that are already out there and being able to read together.

I also had mentioned as what Katie had said about project workflow as another easy behavior creating your system and you can keep it private, as well and then the third behavior was pre-prints. Pre-prints is something I could-- It's easy because it's

something that just involves sharing your material as Nick also mentioned earlier on in the process to get feedback from folks. I think that as a grad student early on, I was pretty anxious about doing that being in the learning process myself.

Putting a pre-print doesn't always have to be before it's accepted, it can also be at different stages to make it more accessible to folks. Earlier on in my career as a grad student, I maybe was more nervous about that and waited for it to be accepted and if it's allowed with the journal, then put it out there so that more people can access the article. Later, as I'm getting more comfortable, there's certain things that I'm more willing to put earlier on and get that feedback which has been really helpful. There no also one-size-fits-all and that's something else.

Then this is not in the article but what Nick was saying, I just wanted to also mention, if seeing a pre-registration, if your advisor is not someone that wants to get put that posted or they're uncomfortable with that, you can still do a pre-registration for yourself and think through all your questions and think through what kind of power analysis and go through the motions and be able to use that in discussion. Even if maybe that second step is not quite there, everyone's not on board for putting it on OSF yet, I think they'll think it's a skill that you can do. We put that as medium but I think that doing it for yourself and still can be part of the process that can be helpful.

Peggy: All right. It's a tool to just organize your thoughts and I liked what Nick said about putting it there so that-- You can be organized and forget things, [chuckles] it's a really good handy tool. With pre-prints, I can see the hesitation initially to put your work out there when it's your personal work, of course, you don't want criticism but good constructive criticism I think helps advance your work, and it's just taking that first step. I know is very hard for most people, myself included. I don't want to exclude myself. Nick, did you have something to say?

Nick: Yes, I was just going to jump in really quick to follow up Ummul's comments. I know that sometimes there's a worry out there about pre-registration and a risk of scooping. The idea that you put your work out there and it's public, and he said, I'm going to study the effect of X on Y and someone goes, "Oh, that's a great idea." I bet I can do it faster than you and get it out there.

At least on the OSF pre-registration, they can be hidden for over two years. You can actually hide that, you can timestamp it for the date that you are planning but then it doesn't actually become public and visible for two years. This risk we're trying to mitigate, to make as minimal as possible that if you are planning on it in 2021 and you can choose when it gets unmasked if that is a concern. I felt that come up a little bit, I just wanted to head off that potential risks that people see.

Peggy: Absolutely. That's good to know. Katie, you mentioned in your book chapter that many Open Science practices boil down to changing your work habits. How easy is it to change these habits? We talked a little bit about this, about your advisor or mid and later career, psychologists may not want to change your habits yet or they didn't learn about open science in their schooling. What do you say to that? How do you help people change their habits so that they're doing Open Science?

Katie: Yes. Great question. To the first question, how easy is it to change your habits? Once your habits get formed, once you get set in your ways, if it comes lots

of other behaviors, these behaviors perpetuate themselves. As a early career person, as a graduate student, I always talk about the importance of establishing good habits early. When you don't have a set workflow or a particular way that you do things, that's the easiest time to get started with new practice and to integrate it into your everyday.

For someone that's a little bit later on in your career, absolutely, you can change your behaviors, change your practices but it is going to take a little bit of work to figure out what works for you. That's really what folks need to do to have these practices is to investigate them and try them out and figure out what works for them and what doesn't, it's a trial and error process.

Peggy: Absolutely. We touched a little bit upon the Open Science Framework but I want to talk more about it because it's obviously a very important research management tool and we got lots of questions about it. Nick, can you just tell us a little bit more about the Open Science Framework? What is it exactly?

Nick: Great question. The Open Science Framework, it's basically a storage management tool. If you can think of it as, those little Russian dolls, they open it up and there's other one inside of it, that's like the OSF, it's very flexible and you can use it to store things however you want. I think that's where a little bit of the hesitation comes from is that it's very flexible and very open and that there's no right way to use it.

The way I use it, I think of each project as a research project where I want to study some effects of X on Y and so I will make a project on the OSF and that gives me a whole bunch of tools, a whole bunch of places I can put things I can make another component, we call them, it's basically a folder. They put a folder and say, "Okay, well, here are all the materials that I need to use to execute that research." Another folder here is the manuscripts in progress, here are the things that I'm throwing back and forth with my collaborators.

You can invite collaborators to your projects, they can see all the different pieces you can control, who sees what piece let's say. You only want your maybe the people who you're writing the manuscript with to see the manuscript portions but not the other portions. It's basically a sandbox of storage that is very secure. We have plans for, "Oh my God, what happens if currency fails?" "Is all your data lost?" No. We have contingencies for all of that. It's a very safe place to store your research, very easily shareable place with lots of controls for viewing and sharing and permissions, and a community of sharing materials.

Peggy: Logistically, you go to the OSF webpage and you have to create an account, is that right?

Nick: Yes, that's right. Let's say you wanted to jump in. As soon as this webinar ends, you want it to jump right in. You would go to osf.io. If you don't have an account, you would go and create an account and it just takes an email and a password. Then from there, you're all set. You can make a project, you could make a pre-registration which now doesn't even need a project but you could either do it before or after you make a project.

You can start putting data, you can link, let's say you have data on Google drive or you have manuscript drafts on Google drive, you can link in third-party tools, Dropbox, Google drive, things like that and you can get started right away. Again, it's very much a sandbox, you can build a big castle or you can just build a mound of stuff. That was how I used to play in the sandbox. It's very easy to jump into and then you make of it what you want to make.

Peggy: That's great. That's good to know. Katie, do you use a project template in your lab to organize your OSF projects and what does something like that look like?

Katie: Yes. I definitely do recommend that people use a project template. You certainly don't have to just get started, but what it does is it gives you a little bit of structure and it gives you a little bit of consistency across different projects so that you don't forget anything. Once you begin adding these different components, these different pieces, you might have your data, your materials, the Manuscript, maybe a video demonstration of someone running through the protocol. There's all these different little pieces you want to make sure that that information is organized not just for you, but for any viewer or reader who's going to be consulting those materials. Having a template is a way to have some consistency both for the reader and for you.

Peggy: It's a way to organize things. Ummul, how would you rank the difficulty level of creating an OSF project? Is it easy, medium, difficult? It depends? What do you say?

Ummul: In the paper, we put it under project workflow so in the easier category, but I think that as Nick had mentioned, you could really build a castle, so you can make this more and more complex for yourself. Include more detail and in creating an OSF page or how you create your structure. As Katie was talking something else popped into my mind of like what was Nick was saying is like how would you incentivize yourself to do some of these things? This also actually will save you a lot of time in the long run not only helping your future self with remembering where things are, if you ever move institutions, you always have access to your OSF page.

You can easily share it with people all over the world and not have to worry about who doesn't have Google but they don't have Google drive or something like that. It's just an easy way to get people on the same page and then, for me, even with now that I'm on the reproducible coding stage, my pathways for different projects. For those of you who do do coding, this might sound familiar, but some of that is consistent. It's just like less work as you keep going [unintelligible 00:32:17] a lot of time.

[crosstalk]

Katie: Sorry, Nick. I was thinking about when it comes time to share data or materials along with a publication, that can also be really easily facilitated through your OSF page. That stability that Nick mentioned is a really important feature of it so I wouldn't want folks to think that OSF is the only tool that you can use for this, but it does have some nice features that make it a good tool.

The things you want to look for when you're thinking about a tool to share supplemental materials with a journal article is you want there to be a stable link, so something like a Digital Object Identifier DOI, but something that's going to be persistent so that later 10, 20 however many years later, when people want to access that material it's still there where you said it would be. The OSF has really nice for that long-term storage because we can have some confidence in the stability of the materials.

Nick: That's a really great point. To back off of that, I used to work in a biology lab where we had to keep a zip drive and all the zip disks because who knows when you needed it. I was just going to jump in really fast and say, Ummul, since you mentioned coding, the OSF also has a very flexible API and we also have an art package called OSF art. You can actually even build in the updating and sharing of materials into your code so if you're looking for an end-to-end workflow that's reproducible from the very beginning up until moving materials into an OSF project, there is capability for that as well. We're really trying to do the best we can in that regard.

Peggy: I really wish the OSF existed when I was in graduate school because I used to inherit data from graduate students who had moved on and graduated, sometimes it was a nightmare; there was no codebook or you had a disc here or there and a bunch of paper and you had to make sense of it. Putting something like that into the OSF, I think would have made my life certainly a lot easier, not just my own work, but work that you may inherit from others in your lab.

Nick, when I spoke with Brian Nosek in the webinar in April, he mentioned that traffic to the OSF webpage has just blown up just exponentially. I suspect this is because of COVID because conferences have been canceled and people are posting there more. Is that true? What have you seen in terms of traffic on the OSF?

Nick: Prior to COVID growth had been steadily exponentially growing as well. Even before up to 2019, there has been very large growth in the usage of open science framework and looking at number of preprints submitted across all the metrics of our tools. There has been exponential growth and in COVID, it has even gone up more. I suspect part of that is because across all these tools and across transparency, the movement of information is much faster.

I know we all went through the past year and day by day we were looking for more information, new information about what was going on, more the virus, more about treatments and preprinting, this goes across different tools as well like Katie mentioned, we're not the only tool in town but we think our tool is pretty good. I have to sell it a little bit. [chuckles]

Peggy: It is good though.

Nick: If you look readership and submission of materials that were COVID specific it was very high because rapid collaboration to learn about this novel virus and about treatments for it was very important. The publication system as it exists now doesn't always facilitate super rapid transformation of transmission of information. I think during that we saw a large usage of our tools and of preprinting, but that was even the case even before COVID

Peggy: Great. Katie, you're Executive Officer for SIPS and last year SIPS conference became completely virtual as many other conferences did. All the presentations and posters for that conference were put into the Open Science framework. How did that process work out for SIPS?

Katie: Really well. We had discussed having virtual components to our meetings prior to 2020, but COVID obviously really forced the issue. I think we saw a much greater accessibility to the conference for people that otherwise wouldn't have been able to attend. We're getting ready to have our next meeting in just about two and a half, three weeks, something like that. We'll also be fully online and it just seems like our future meetings will always include some hybrid or online component to increase that accessibility.

Open sharing of the conference materials has also spread their reach so all of the past conferences are still live on the web so you can go to improvingpsych.org/sips 2017, '18, '19, whatever. They're all still there and all the materials from all of those meetings are available. There's a lot of shared resources that people have created on all kinds of Open Science related topics.

Peggy: I think it's great to have that information available online because, in the old days of conferences, you had to write notes for presentations or take pictures of the posters. Now you have all of this available online which makes it so much easier to remember what you saw and heard. I wanted to shift our attention a little bit more now to pre-registration because we got lots of questions about that. Nick, you defined what pre-registration is and why it's important, but there seems to be a great deal of apprehension and uneasiness when it comes to this in particular. Why do you think that's the case?

Nick: That's a big question. I think there could be lots of reasons. I can give you candid personal experience. When I did my very first preregistration, it was extremely long and extremely detailed and I wasn't expecting it. I went, oh, I'll go to OSF. **[unintelligible 00:39:19]** this was when I was a graduate student. I'll go there, I'll do it. They had a preregistration challenge at the time, which meant if I did a good job, maybe I would get some money.

That again, the selfish **[unintelligible 00:39:29]**. It was very daunting and it was only until going through it did I realize, okay, these are actually important steps. Every step was important and really had to think about it before filling it out. I went in, oh, I'll fill this out. Then I'll start my project. It's something really got to devote a little bit of time to conceptually thinking through your project from beginning to end and unfortunately, it's sad to say it, but that was a little bit of a change, because when you come in to graduate school, it's do the work, get the publications, write, write, write, work, work, work. A little bit of conceptualizing the process, which was great. Obviously, a breath of fresh air getting towards that scientific method ideal. I think maybe just this idea that there is a burden to it, that it might be difficult, is maybe one reason why there's a little bit of apprehension or a little bit of caution towards it.

Another one that I mentioned before might be the idea that I'm going to put this out there before I do the work, and then maybe someone might take that idea and do it faster or better, or in a way that gets published easier than my work. That's a concern that I've heard a bit. I haven't really seen any evidence towards it being a

problem towards it happening. I'm sure there's anecdotal cases out there, but I haven't seen any evidence of it being a problem.

Again, there is you can timestamp it, and block it for two years on the OSF. There are also other ways to pre-register as well outside of the OSF. A colleague at the school I went to for graduate school, Rutgers, they would pre-register using Qualtrics, so they would make a Qualtrics survey, but instead of putting in questions, they would put in all of the details of their project, and then launch it, and that would give them a timestamp, so they can say, "Look, if you have any questions you think I didn't plan this, here is the Qualtrics at the timestamp of when it was launched, and this is what I planned to do ahead of time."

There are many different ways to do it. I think sometimes, seeing a long template can be a little discouraging, but there are many ways to enter that pool. There's a very shallow end with just an open-ended, "Here's my hypothesis," and then everything else is still left open, up to, "Here is the equation for the regression that I'm going to run. Here is the test that I did on a 5% data holdout," all the way through to as elaborate as you want to get.

Peggy: [crosstalk]. Go ahead please, go ahead, Ummul.

Ummul: I'm sorry, I was just going to reiterate the last point that Nick made, because I think that's kind of sometimes people can be anxious about doing a pre-registration because they think they need to be, the only way that it counts if you write the regression equation for the thing that you're going to test, but really putting your research questions and hypotheses like expectations, that is a pre-registration.

That's already moving in the right direction, and that doesn't mean that if you want to have your data and you see some other interesting associations that you want to share, or you want to be able to look more into, you can do that. You just say that this wasn't pre-registered, this is exploratory, or this was added on, and you're just transparent and clear about that. I think that sometimes people worry that, too, that you can't then look at other things.

Nick: One little brain worm that I try to say it every time we talk about it, is that pre-registration is a plan and not a prison. Think about coronavirus, we had lots of projects that were pre-registered, they were going to collect data in person. They're going to have students come to the lab, and then March 2020 came around, and everything had to shift, so we had to change, and things went online, and plans changed.

That is totally fine, because then you just say, "Okay, well, we pre-registered this at a certain time with a certain expectation of how things are going to go, to the best of our ability. We did this type of change. We moved it online and so on, and so on." Like Ummul said, there is so much more to explore, and knowing that you're exploring and generating hypothesis, and not trying to confirm hypothesis is a really great distinction. Knowing what you're trying to confer, a priori hypothesis, and what you're exploring, and what you're generating, having that distinction clear in a manuscript or clear in a project, I think is really valuable.

Katie: You guys have brought up a lot of points that are things that I hear all the time from folks when they're thinking about pre-registration for the first time, there's definitely that apprehension about getting it wrong, right? That if I freeze this plan in time, and I make a mistake, that this will be some big ordeal to do something else. As you all have been saying clearly, the best thing to do is to just admit when the plan needed to change, and explain the reason for the change transparently.

Another really important distinction that I like to emphasize when people are just getting started with this is the distinction between pre-registering a design, an analysis plan, and your hypothesis. Sometimes when people talk about pre-registration, they could be using the word pre-registration and talking about one of those things, but not the other, and someone else can be coming at it from the opposite direction. A lot of conflict and confusion can be avoided if you're clear about which elements of your project you're focusing on pre-registering.

If you're trying to do a test of a theory, and you want to have a chance to falsify a hypothesis, that's a slightly distinct goal from trying to fix certain parts of your analysis plan so that you don't over-fit your model to your data. That's different from saying we constrain the details of the design of our study, and we can constrain those details of design, but not necessarily even have hypothesis or not necessarily even me doing a quantitative analysis. These are three independent pieces, and being clear about which one of those you're working on or you're trying to tackle, is really helpful, both for you and for your audience.

Peggy: We're getting questions from the live audience that they're still concerned about registering their work and others stealing it. I just want to reiterate what Nick is saying that you don't have to make it public. This is a big concern. People are hesitant because they are afraid that others will claim their work, but just let's reiterate, you don't have to make it public, you said for what, two years?

Nick: You don't have to make it public. On OSF, we think about pre-registration on the tool, OSF, we can guard it, we can block it, make it private, for up to two years. You can change that. You can make it one year. You set the date, and the max date is two years from when you're doing it. You can say, "Oh, I want it on Christmas day. I want it to be a nice present to myself."

A pre-registration as a concept, you can do it privately, and you can keep it there indefinitely. The public aspect is less about making it a pre-registration, and it's more about this concept of openness, transparency and sharing. As a concept, the pre-registration doesn't need to have a public element. It's just about before executing your research, testing the hypothesis, or collecting the data, that you plan it out beforehand. Part of that is the timestamp saying, "Look, I did this on January 1st. I laid this all out, and I'm going to start on January 15th." The public aspect is separate.

Katie: I do want to clarify that at journals that issue badges for pre-registration, the requirements are going to be a little bit more stringent. Also, at journals that are looking to review a pre-registration alongside your article, there's going to need to be some kind of accessible link. One of the nice features about OSF is that it has what's called view-only links, which means that you can share the pre-registration with your

identity obscured if you want, for the purposes of masked review, with peer reviewers and with the editor, without actually sharing it more broadly.

Potentially, the pre-registration could stay masked up until the point of publication. I often talk about this option to do private pre-registration as you're just getting started, but eventually, when it comes time to get credit for the practice or to talk about that practice in the context of publication, you're going to want to be using a more formal system.

Some people like OSF There's another system that's quite simple called AsPredicted. I generally tend to not recommend AsPredicted, because it is not an open and searchable repository, so it doesn't have some of the nice features, nice registry I should say. OSF and ClinicalTrials.gov are surgical registries that the scientific community can actually look at and see what's been registered once those embargo periods are over.

Peggy: We have a couple of people submitting questions live, who are asking about the project templates that you have used in the past. This can be for any one of you. Did you create those from scratch or did you use some online resources, or some template that helped you get started online and then you made it your own?

Katie: Yes, for the template that I use in my lab, I have it linked in my chapter that was shared earlier in the chat. You can find that information there, and when I teach workshops on this topic, that's the template that I use to demonstrate, but it's really just file structure that resembles an article. There's a section for materials, there's a section for data. It's just including those different elements to make sure that all of the pieces that you want to include are there in your project. Do you all know of any other online resources for templates?

Ummul: I've mainly used the ones on OSF or the ones that maybe has come out of SIPS at different points, but I've always been-- I've done some mixed method work where I pre-registered. That template looked, or what I use looked a little bit different than maybe a more quantitative study. Then I've always deleted sections that just were not relevant to me, or were just not applicable, and felt free to do that. That's, I guess what I've done so far.

Katie: Ummul, your comments are making me realize that we're using template in two different ways. There's templates for OSF projects, and then there's also templates for pre-registration. The pre-registration templates are also extremely useful and important to use. Also in the chapter, I have a table in there of pre-registration templates that are applicable for different fields. There's something like 25 entries in that table. There are general templates, but there's also very specific ones for things like EEG research, or secondary data analysis, or all different kinds of communities that have adapted pre-registration protocols to their specific field.

Peggy: Yes. Thank you for pointing that out, that there is a distinction. There are pre-registration templates, and there are OSF templates, project templates, and your chapter has examples of both. We're getting some questions from the audience about the open science framework. I'm not sure. Let's see if we can try and clarify this one. Someone is asking, is OSF a global platform for researchers? That's all I got. I don't know. Yes, right? Yes.

Nick: Yes. As far as I can understand, yes, it should be available everywhere if you have internet connection. I mean, it's based in the United States. It's based in Charlottesville, Virginia as a nonprofit, but yes, it should be globally accessible. We're very much open to global collaboration and working, people working wherever they are in the world. We would love to have them be a part and use the product for sure.

Katie: One of the nice features is that you can store your data that is stored within the OSF in a particular country or region. For people from Europe in particular, this can be useful when there are requirements to store data within the local country, rather than abroad.

Peggy: I have another question from the audience about entering data into the OSF. This is a slightly lengthy question. Can you enter previously completed research for future reference? What about older data for which one has already outlined storage plans with institutional ethics boards? Do researchers typically need to obtain further ethics clearance to put it in older data since it wouldn't have been part of the original data storage plans or is this just a 'by institution' issue?

Katie: I think it really depends on the specifics of your project. There's a really nice chapter in the journal AMPPS, Advances in Methods and Practices in Psychological Science, that talks about the practicalities of the ethics of data sharing. It's by Michelle Meyer. Basically, when it comes to existing data we follow the same ethical principles that applied when you got your original ethics clearance.

If this was particularly sensitive data that you said you would store and protect in a certain way, then you would need to either get new permission from your ethics board to share it, or you would need to make it deidentified or apply certain practices to make it so that it's ethical to share after the fact. It's certainly easier to share data if you have obtained the permissions ahead of time, but just because you didn't get them before doesn't necessarily mean that you can't get them retroactively.

For certain types of data, say your standard social or cognitive research study where the participants are anonymous and the data's not sensitive, most of the time it's possible to share that kind of data even while after the fact. It usually involves going back to your ethics board and revisiting the issue of whether or not it's okay to share.

Peggy: Well, speaking of data sharing, we got lots of questions about that, and maybe we can spend the last couple of minutes talking about it. Our listeners are still worried about putting data they've collected into the hands of others for a number of reasons. One question, is there a dark side to the open science movement where sometimes openness can devolve into bullying and data theft? Have you experienced any of this or heard of it personally where people, researchers might get angry because data's not being released quick enough or for whatever reason and there's public criticism for it? Anyone? I've heard about this myself.

Katie: I feel like I've been talking a lot so I want to let other people talk if someone's got anything to say.

Peggy: Well, is this a genuine concern? Should people be worried about this potential dark side to the open science movement?

Nick: I can see it being a concern. I can see where the thought comes from. I haven't seen any bullying from that angle, from the angle of this isn't being shared fast enough, or you're not sharing it the right way. I haven't seen anything like that. Twitter can be a mess, and I would hesitate anyone who goes to Twitter, to take the whole thing with a big grain of salt. I have seen people get into arguments on Twitter, but never about sharing substance. If it's about sharing data, I've only really ever seen just maybe clarifying questions, like, "Oh, I see that you shared the data, do you happen to have a code book that would go with this?" Things like that.

One big positive that I see in sharing data as opposed to a negative is that, a lot of the times in manuscripts, we put materials that data available upon request, which is just asking for more emails, right? I know that, especially if you look on Twitter, emails for academics tend to always be this avalanche, this is Sisyphean stone to push up the hill. A great thing about sharing the data is that you can put the link right there in the manuscript, and it's up to the person who wants to find it to go find it. It's there, and you're now pushing it instead of asking to pull. You're not pulling, you're pushing your data out.

I like that because then it's, okay, the manuscript is there, the data is there. Maybe you'll get a question again, "This is great. Do you have a code book? I don't understand how you named your variables," and either you do or you don't, or you don't get an email reply, and all of those are fine. In terms of bullying or a negative side to sharing, I personally haven't seen that. I've really seen a lot of support actually.

Peggy: Yes. I think there's more of a plus side than a negative to all of this. I think you mentioned, we spoke yesterday, that you can put permissions or restrictions on how much you share data. Is that right? Can you guys talk a little bit about that?

Nick: That's right. Again, like Katie mentioned with the pre-registrations, there's also, you can have a private project for your data and then provide a view-only link. If you don't want to have it out there on the manuscript, you can have it on a private project. If someone asks, you can just give them a view-only link. Again, since it's like a sandbox, you can even say, "Oh, here's the sensitive material that can't be shared. Here's the material that can be shared," and you can determine at what level you feel comfortable and then even provide view-only links for those as well.

There are lots of different ways from completely private to completely open, that you can control that level of sharing. Again, a lot of it comes to ethics, and ability to share. That's also another thing. There are plenty of research fields where the open sharing of data doesn't make sense and that's okay, too. There are sensitive populations, hidden populations, where sharing could be detrimental. Small populations where de-identification may not be possible. That is also okay. I think a push towards openness doesn't mean trampling on those populations, or on those research fields that require more privacy.

Katie: One of the really positive developments that I'm seeing is that journals are beginning to change their policies to require a statement of data availability at a minimum, and then in some cases, actually to require sharing of the data or explanation of the ethical or legal restriction, why it can't be shared. Those policies

are a positive development, because what it does is it shifts the conversation about the availability of the data to the time of publication, as opposed to afterwards.

What happens now is if a person wants to get access to particular data, say to re-analyze it and to establish a claim that another researcher made, there's no guarantee that data will ever be provided. For journals asking authors to give an indication of the data availability ahead of time, it levels the playing field, and makes it very clear to everyone exactly under what conditions the data can be accessed.

That's coming out of APA journals in particular, under a new policy that's rolling out through the rest of this year, following the TOP guidelines, so Transparency and Openness Promotion guidelines. All of the APA core titles, by the end of 2021, will require a data availability statement at a minimum. Some of those journals will also require sharing or an explanation of the legal or ethical reason why the data can't be shared in the manuscript. That just makes everything super clear for authors and for readers.

Peggy: Thank you for making that point, Katie, I'm looking at the clock and the hour has already gone by. We could talk more about this for hours, but unfortunately, we are out of time. I want to thank our guests, Ummul-Kiram Kathawalla, Katie Corker, and Nick Fox for taking the time to speak with us today. It was really informative. I hope you were able to answer many of the questions that our listeners had, and if not, we'll just have to come back another time and do another webinar, but with that, that concludes our webinar. Many thanks to everyone for being here today. Have a wonderful day, and goodbye.

Katie: Thanks, Peggy.

Peggy: Bye, everyone.

Nick: Thank you. Thanks everyone.

[01:01:50] [END OF AUDIO]