Evaluating the pedagogical effectiveness of study preregistration in the undergraduate dissertation: A Registered Report

Madeleine Pownall[1*]

Charlotte R. Pennington[2]

Emma Norris[3]

Kait Clark[4]

1. School of Psychology, University of Leeds, UK
2. School of Psychology, Aston University, Birmingham, UK.
3. Department of Health Sciences, Brunel University, London, UK.
4. Department of Social Sciences, University of the West of England, Bristol, UK.

***Corresponding Author:** Madeleine Pownall, School of Psychology, University of Leeds, UK. E-mail: M.V.Pownall@leeds.ac.uk

Article Type: Registered Report

Submission: PCI Registered Reports, with final submission to Royal Society Open Science

(Level 2 taxonomy)

CRediT Statement

Conceptualisation: MP, CRP, EN, KC; Data curation: MP; Formal Analysis: MP, KC; Funding acquisition: MP, CRP, EN, KC; Investigation: MP, CRP, EN, KC; Methodology: MP, CRP, EN, KC; Project administration: MP; Resources: MP; Software: MP; Validation: MP, CRP, EN, KC; Visualisation: MP; Writing – original draft: MP, CRP, EN, KC. Writing – review & editing: MP, CRP, EN, KC.

Abstract

Research shows that questionable research practices (QRPs) are present in undergraduate final-year dissertation projects. One entry-level open science practice proposed to mitigate QRPs is 'study preregistration', through which researchers outline their research questions, design, method and analysis plans prior to data collection and/or analysis. To date, no research has examined the effectiveness of preregistration on undergraduate students' learning and perceptions of research practices, despite recent recommendations that preregistration could facilitate engagement and reduce anxiety with the dissertation process. In this study, we aim to empirically test the utility of preregistration as a pedagogic tool in undergraduate dissertations. A total of 200 UK psychology students will be recruited and classified into two groups: those who preregister their empirical quantitative dissertation (n = 100; experimental group) and those who do not (n = 100; control group). Attitudes towards statistics and QRPs and understanding of open science practices will be measured both pre- and post-dissertation. Exploratory measures include participant's capability, opportunity and motivation (COM-B) to engage with preregistration, measured at Time 1 only. In line with/contrary to hypotheses, study preregistration [significantly/did not significantly increased/reduced] positive attitudes towards statistics, endorsement of QRPs, and understanding of open science. Exploratory analyses indicate that preregistration was associated with [greater/less/no difference] capability, opportunity and motivation and qualitative responses revealed that preregistration [XXX]. These results contribute to timely discussions surrounding the utility of embedding open science principles into undergraduate training.

Keywords: Preregistration, open science, reproducibility, undergraduate training, dissertations; research training

Evaluating the pedagogical effectiveness of study preregistration in the undergraduate dissertation: A Registered Report

In recent years, psychology has put reproducibility, replicability, and transparency at the forefront of the research agenda (Asendorpf et al., 2013; Munafò et al., 2017; Open Science Collaboration, 2015). Fuelled by replication concerns in the general scientific literature, an era of 'Open Science' has prompted a plethora of ideas and recommendations to envision a new future for science (Pashler & Wagenmakers, 2012). A move to study preregistration, open materials, and open data are proposed to combat *questionable research practices* (QRPs; John et al., 2012) that plague the literature, such as *p*-hacking (Head et al., 2015), Hypothesising After Results are Known (HARKing; Kerr, 1998), and selective reporting (John et al., 2012) or 'undisclosed flexibility' (Simmons et al., 2011). Furthermore, an incentive shift to high-quality, *slow* science is picking up momentum (Frith, 2020). Despite these practices being endorsed and embraced by the scientific community, scant research assesses the pedagogic value of Open Science practices in improving teaching and learning. In this study, we examine the value of study preregistration in the undergraduate curriculum to assess whether this can improve attitudes towards statistics and QRPs, as well as understanding of Open Science. Study preregistration comprises a time-stamped, uneditable protocol that transparently outlines a study's research questions, design, hypotheses, methods and analysis plan prior to data collection and/or analysis (Nosek et al., 2018; van't Veer & Giner-Sorolla, 2016). The process of preregistration encourages researchers to plan the decisions that have traditionally been made after data collection (e.g., exclusion criteria, analysis details) using a wide host of platforms such as the Open Science Framework (https://osf.io/) and AsPredicted (https://aspredicted.org/). Preregistration increases transparency about the authors' original intentions (LeBel & Peters, 2011) and should, in theory, limit selective reporting of results (Nuzzo, 2015).

Deleted: compat

Importantly, much of the recent shift to Open Science practices has been championed by grassroots, collaborative initiatives (e.g., see Button et al., 2020; Pownall, 2020b). In recent years psychologists have developed initiatives such as the Society for the Improvement of Psychological Science (SIPS; https://improvingpsych.org), the open source reporting forum PsychDisclosure (LeBel et al., 2013), and the early career researcher-led journal club, ReproducibiliTea (Orben, 2019), all with the aim of improving the rigour and reproducibility of psychological science. Beyond these, organisations and initiatives are centred around the improvement of psychological science, stressing the importance of rigorous, robust methods (e.g., Crüwell et al., 2019; Munafò et al., 2017; Simmons et al., 2011; Tennant et al., 2016; Wagenmakers et al., 2012). For example, Klein et al. (2018) note the importance of preparing and sharing research in a way that values transparency and note how this can be done incrementally to improve research efficiency and credibility. Similarly, Devezer et al. (2020) focus on recommendations to improve methodological problems in science reform, such as the adoption of a formal approach that embeds statistical rigour and nuance into science reform.

Open science in undergraduate training

These recent shifts towards novel and creative ways of promoting uptake of Open Science practices offer the opportunity to reevaluate core aspects of undergraduate training, as well as wider scientific research practices. Specifically, the Open Science movement invites a reconsideration of research training to address some of the ongoing pedagogic problems that exist in undergraduate teaching. For example, there have been some emergent initiatives that have specifically concentrated on how to embed teaching on the 'Replication Crisis' and Open Science practices into undergraduate teaching (e.g., Button et al., 2016, 2018; Chopik et al., 2018; Frank & Saxe, 2012; Janz, 2016). There has also been a keen interest in interventions to improve understanding of QRPs in, for example, graduate psychology training (Sacco &

Brown, 2019; Sarafoglou et al., 2020). However, the impact that these have on students' learning and perceptions is yet to be empirically investigated.

The undergraduate dissertation

Final-year psychology dissertations typically consist of an independent empirical project that requires students to design a protocol, collect data, and analyse the results. According to the accreditation standards of the British Psychological Society (2019) undergraduate psychology dissertations in the UK require students to "individually demonstrate a range of research skills including planning, considering and resolving ethical issues, analysis and dissemination of findings" (p. 13). Final-year projects are thus typically self-contained research studies that are constrained by the scope and availability of resources but are supervised closely by an experienced academic. Much pedagogic research has demonstrated that, given the level of autonomy that students have over their final-year dissertation, students typically struggle with some of the components of this mandatory part of their degree. For example, it is reported widely that undergraduate students face anxiety, disengagement, and stress related to their final year dissertation (e.g., Devonport & Lane, 2006). Indeed, research shows that undergraduate students often experience difficulty with their dissertation, due to pedagogic issues such as debilitating statistics anxiety (e.g., Onwuegbuzie & Wilson, 2003), underconfidence with their writing ability (Greenbank et al., 2008) and challenges navigating supervisory relationships (Day & Bobeva, 2007).

Contemporary research also indicates that QRPs are prevalent within undergraduate research projects (Krishna & Peter, 2018; Kvetnaya et al., 2019; Sorokowski et al., 2019). For example, Krishna and Peter (2018) assessed the prevalence of QRPs in final-year undergraduate dissertations and found that students typically engage in QRPs related to reporting and analysing their results. Similarly, Olson et al. (2019) studied the prevalence of questionable research practices of taught masters students' theses and found inconsistency of

p-value reporting, although this was not clearly a result of intentional *p*-hacking. Research outside of psychology also indicates that from dissertation to publication, the ratio of supported to unsupported hypotheses more than doubles (O'Boyle et al., 2017). Recently, there has also been a focus on addressing QRPs that feature in undergraduate final-year projects through consortia based approaches (Button et al., 2020; Kvetnaya et al., 2019; Munafò et al., 2017) and through focusing on replication studies with undergraduate projects (e.g., de Leeuw et al., 2019; Jekel et al., 2020).

The use of QRPs in the undergraduate dissertation likely stems from many different sources: resource and time constraints mean that many undergraduate experiments are typically underpowered (Button et al., 2016; 2018), students perceive that there is a pressure from supervisions to 'find' significant results, which are more likely to lead to a publication (Wagge et al., 2019), and in our own experience, report worries that a 'lack of significant' results will adversely affect their grades. QRPs may also stem from a lack of awareness that they are problematic (e.g., Banks et al., 2016). This mirrors the pressures put on academics to publish novel, positive results (Franco et al., 2014), due to the 'publish or perish' culture that pervades academia (Grimes et al., 2018). If these studies are then selectively published, they contaminate the scientific literature with unreliable results. Indeed, an undergraduate publication is seen as an advantage when applying for highly competitive places on taught masters and doctoral training (Button, 2018). Understanding undergraduate students' use and acceptance of QRPs is useful, given that students' research behaviour reflects the quality of Open Science teaching and adoption of rigorous practices more broadly (Olson et al., 2019). Some emergent research has begun to investigate the research practices of early-career researchers (Nicholas et al., 2017), including uptake of open science practices (Stürmer et al., 2017).

Importantly, consideration of the prevalence of QRPs in the undergraduate dissertation has led to interventions to reduce them. Button et al. (2020), for example, describe and evaluate

Deleted: published

an approach to improving rigour of undergraduate dissertations via a consortium approach to science. This approach also echoes Detweiler-Bedell and Detweiler-Bedell's (2019) team-based approach to undergraduate research supervision. Creaven et al. (2021) stress the importance of embedding a concern for rigour, transparency, and openness into the undergraduate dissertation, stressing how the undergraduate dissertation should be thought of as an important learning activity that offers many pedagogical benefits to students. Similarly, Blincoe and Buchert (2020) propose that preregistration may be a useful pedagogical tool for undergraduate psychology students. Despite some useful and recent conversations that discuss the *need* to embed an open science approach into undergraduate research training (Button et al., 2020; Creaven et al., 2021; Pownall, 2020), an *empirical* exploration into how Open Science practices in undergraduate dissertations may benefit (a) students, and (b) the open science movement has been notably absent from these conversations. Indeed, while much work has considered how to promote uptake of preregistration practices of early career (Zečević et al., 2020) and more established researchers (Kidwell et al., 2016; Munafò et al., 2017), little research has explicitly focussed on the utility of preregistration for undergraduate students' research practices, despite recommendations that preregistration could facilitate engagement with the dissertation process (e.g. Nosek et al, 2018; Pownall, 2020a).

Commented [KM1]: Most of this preregistration focuses on (a), and it would be good to add more content related to (b) – see my additional comment at the end of the Intro section.

The value of preregistration

Preregistration is one entry-level way of establishing a level of rigour and robustness into the undergraduate dissertation process. The potential value of preregistration in this context has been noted by educators too. For example, the Framework of Open and Reproducible Research Training (FORRT; www.forrt.org) includes preregistration as one of the six pillars of effective reproducibility training, including at the undergraduate level. Others have suggested that "most study programmes should offer easy ways of implementing preregistration in empirical research seminars" (Olson et al., 2019; p 13). As Pownall (2020a)

also argues, the process of embedding preregistration of undergraduate dissertations largely complements current practices in dissertation supervision. Sacco and Brown (2019) note that preregistration is thus useful when conducting research with the view to publish the results with undergraduate students (see also Blincoe & Buchert, 2020).

The present study

We aim to empirically investigate the utility of preregistration in undergraduate dissertation provision. Our core research questions aim to evaluate whether preregistration is a useful pedagogic practice to improve students' attitudes towards statistics, endorsement of QRPs, and understanding of open science in this cohort. To achieve this, we will employ a 2 (Group: preregistration vs. control) x 2 (Time: time 1 pre-dissertation vs. time 2: post-dissertation) mixed design. We have three confirmatory hypotheses based on a significant two-way interaction between Group and Time:

H1: Students who preregister their dissertation will have higher positive affect towards statistics, higher self-reported competence with statistics, higher perceived value of statistics, and less difficulty with statistics at T2 compared to students who do not preregister their dissertation.

H2: Students who preregister their undergraduate dissertations will have a reduced endorsement of QRPs compared with students who do not preregister their dissertation.

H3: Students who preregister their undergraduate dissertations will have higher confidence in Open Science terminology compared with students who do not preregister their dissertation.

As an exploratory measure from a behaviour change approach (Norris & O'Connor, 2019), we will also assess students' Capability, Opportunity and Motivation (COM-B) towards preregistration at Time 1 only, as well as qualitative responses regarding the perceived value of preregistration at Time 2.

Commented [KM2]: One thing missing from the Introduction section is the predicted impact the results of this study will have on the open science movement. How will your findings help facilitate or change methods and understanding of open science?

Method

Transparency Statement

All materials and data will be publicly available via the Open Science Framework: https://osf.io/8sndy/ and our study meets Level 6 of the PCI RR bias control (https://rr.peercommunityin.org/help/guide_for_authors). In the sections that follow, we report all measures, manipulations, and exclusions.

Design & Participants

The study comprises a 2 (Group: preregistration vs. control) x 2 (Time: pre-dissertation vs. post-completion) mixed factors design, with Group as the between-participants and Time as the within-participants factor. To be eligible for inclusion, participants are required to confirm that they are a final-year undergraduate student, studying Psychology at a U.K institution. This is to ensure that the study can contribute directly to existing pedagogic policy discussions regarding embedding Open Sciences within the undergraduate dissertation (e.g. the British Psychological Society's course accreditation guidelines).

Our planned sample size is based upon resource and time considerations including the time window for participant recruitment and available funds for participant remuneration (see Lakens, 2021). Two-hundred and forty final-year undergraduate Psychology students will be initially recruited with approximately 20% attrition expected at Time 2 based on prior research sampling from online platforms (Palan & Schitter, 2018). The final planned sample size is therefore 200 participants (*M*age = , *SD* = , % female), with an experimental group of approximately 100 having initiated a preregistration of their final year quantitative project and a control group of 100 not initiating a preregistration. Simulation based power analyses conducted using the superpower shiny package (Lakens & Caldwell, 2021) with 10,000

simulations indicate that this sample size will have 80% statistical power to detect an effect size of *np2* = .04 for the two-way interaction between Group and Time, and 80% power to detect small-moderate effects of *d* = .40 for the focal pairwise comparison between preregistration vs. control at Time 2 (Code/Output: https://osf.io/y9vz7/). Sensitivity power analyses conducted on the final sample size indicate that we achieved XX power to detect effects of > .XX, which was higher/lower than planned. All participants will provide informed consent. Ethical approval has been granted from the local School of Psychology Ethics Committee on 8th July 2021 (Reference: PSYC-266; https://osf.io/5rtch/).

Commented [KM3]: Can you define what np2 is?

Recruitment Plan

We will purposefully sample students via Prolific Academic, university participant pools (SONA) and through social media adverts, ensuring they meet the inclusion criteria. After reading a brief definition of preregistration, participants will be asked to confirm at Time 1 and 2 whether they have preregistered their undergraduate dissertation or not. Because preregistration is typically at the supervisor's discretion, and not widely implemented within undergraduate degree programmes, we will also engage in targeted recruitment to the preregistration condition through appropriate Open Science teaching channels: these include organisational stakeholders such as the UK Reproducibility Network and the BPS, as well as UK institutions who incorporate preregistration as part of their undergraduate curriculum (see Table 1). All participants recruited via Prolific Academic will be paid the equivalent of £6.50 per hour for their time.

Table 1. Potential universities for student recruitment who offer preregistration within the final-year curriculum.

University	**Preregistration approach**
Bath Spa University	Students complete an internal preregistration in Semester 1.
University of Glasgow	Open Science forms an integral part of core undergraduate teaching.
Royal Holloway University	Internal preregistration is embedded into dissertation supervision.
Durham University	Internal detailed preregistration forms part of the final-year dissertation module.
Undergraduate consortium approach (see Button et al., 2016)	The Universities of Bath, Cardiff, Aston, Liverpool and Bristol run group consortia projects through which students' preregister their final year project.

Procedure

Data will be collected online using Qualtrics (https://www.qualtrics.com/uk/). At Time 1, participants will be enrolled for their final year but will not have initiated their dissertation project (September - October 2021). This will provide a baseline in which to compare responses at Time 2 (post-dissertation; May-June 2022).

Participants will first provide demographic information (age, gender, ethnicity, institution of study) before confirming that they are in the final year of their BSc undergraduate psychology degree and plan to undertake a quantitative dissertation project in the 2021-2022

year ("yes/no"). Participants who answer 'no' will be informed that they do not meet the inclusion criteria for the study. We will then collect data related to students' academic attainment in the mandatory statistics module of their degree (scored on a categorical scale from 80-100% to < 40%) and their average grade in the second/penultimate year of their degree. This is to control for potential baseline differences between our two groups.

Participants will then be provided with a brief definition of preregistration, adapted from Lindsay et al. (2016): "*Preregistering a research project involves creating a record of your study plans before you look at the data. The plan is date-stamped and uneditable. The main purpose of preregistration is to make clear which hypotheses and analyses were decided on before you collected your data and which were more exploratory and driven by the data."* Then, to ensure participants have not yet preregistered their project at Time 1, we will ask participants whether they *plan* to preregister their undergraduate dissertation (yes/no) and whether the undergrad dissertation has already been preregistered (yes/no).

Measures (Time 1)

Survey of Attitudes Toward Statistics (SATS-28). To assess whether preregistration improves attitudes towards statistics, students will complete the Survey of Attitudes Toward Statistics (SATS-28). This 28-item scale includes items related to statistics affect (e.g. "I am scared by statistics"), cognitive competence (e.g. "I can learn statistics."), value (e.g. "Statistics is worthless") and difficulty (e.g. "Statistics is highly technical"). These items are scored on a 1 *(Completely disagree)* to 7 *(Completely agree)* Likert Scale and 19 items are reverse scored. A total score will be computed for each of the subscales of the Survey of Attitudes Toward Statistics (SATS-28): statistics affect, cognitive competence, value, and difficulty. Reverse scored items will be re-coded so that higher scores indicate: more positive affect, higher competence, higher value and lower difficulty. This scale has been found to have acceptable internal reliability (Cronbach α 0.64-0.85 for each of the subscales; Dauphinee et al., 1997) and

Commented [KM4]: The categories are not clear

Deleted: this

Commented [KM5]: It would be good to have all measures from the surveys listed in supplementary material

Deleted:

for the scale as an overall index (a = 0.91; Ayebo et al., 2020). The internal reliability of each subscale was poor/adequate/excellent (*Cronbach's a,* affect = XX, competency = XX, value = XX, difficulty, = XX) in the current study.

Attitudes towards QRPs. To assess whether preregistration influences attitudes towards QRPs, students will rate their views on 15 research decisions (11 of which are QRPs, 4 of which are neutral/acceptable) on a sliding scale from 1 (Sensible) to 7 (Problematic; Krishna & Peter, 2018). These include items such as "selectively reporting studies" and "deciding to exclude data after looking at results" (QRPs) and "reporting effect sizes" (neutral/acceptable). The 'neutral/acceptable' items will not be analysed but will mask the nature of this questionnaire. We will compute all 11 items pertaining to QRPs into one total indicating general acceptance of QRPs, where higher scores indicate less acceptance of QRPs. The internal reliability of this questionnaire was *poor/adequate/excellent* (a = XX) in the current study.

Understanding of Open Science. To test current understanding about Open Science practices and terminology, students will indicate their confidence in their ability to define 12 key terms (e.g. Replication Crisis, *p*-hacking, open data, file drawer effect) on a 1 (Not at all confident) to 7 (Entirely confident) Likert scale. These concept recall items will be compiled into a total score of Open Science understanding. The internal reliability of this questionnaire was *poor/adequate/excellent* (Cronbach's a = XX) in the current study.

Attention check. As an attention check, we will add an item "*Please select agree to this question*" in the COM-B measure (see below), to assure data quality. This will be repeated in Time 1 and Time 2. Any participant who fails the attention check will be excluded from the final analyses.

Exploratory Measures

Capability, Opportunity and Motivation (COM-B) towards preregistration. In line with Norris and O'Connor (2019), we will also apply a behaviour change approach to assess the facilitators and barriers to study preregistration at Time 1 only. The COM-B model (Michie et al., 2011) posits that a behaviour occurs only if an individual has sufficient Capability, Opportunity and Motivation to perform it. Capability includes psychological capability (i.e., knowing how to perform the behaviour) and physical capability (i.e., being physically able to perform the behaviour). Opportunity includes social opportunity (i.e., being around others who are performing the behaviour) and physical opportunity (i.e., having the time and resources to perform the behaviour). Motivation includes reflective motivation (i.e., plans and beliefs to perform the behaviour) and automatic motivation (i.e., desires, impulses and inhibitions towards the behaviour; Michie et al., 2011). The brief measure of COM-B developed by Keyworth et al. (2020) will be employed. This measure contains 6 items, where one item addresses each of the six components of the COM-B on a 11-point Likert scale ranging from 0 (Strongly disagree) to 10 (Strongly agree). Each item is accompanied by an explanation of what the COM-B component referred to in the questions means. For example, '*I have the PHYSICAL opportunity to preregister my undergraduate dissertation*' is accompanied by the explanation defined by Keyworth et al. (2020) '*What is PHYSICAL opportunity? The environment provides the opportunity to engage in the activity concerned (e.g sufficient time, the necessary materials, reminders)*'. A total score will be computed for each subscale. The internal reliability of these items was *poor/adequate/excellent* (Cronbach's α = XX) in the current study.

Post-dissertation (Time 2)

The same sample of students will be asked to complete the above measures again at Time 2, which represents a follow-up after their dissertation is complete in May 2022. In addition, participants will also be asked four questions assessing whether they have implemented other open science practices associated with their dissertation: (1) creating an

Open Science Framework account, (2) uploaded material (*open material*), (3) code/scripts (*open code*), and (4) data (*open data*) to a public archive. This will be used descriptively to gain more insight into other contextual factors that are associated with preregistration. Qualitative responses of students' experiences of the preregistration process, including enablers and barriers, will also be collected through three-open ended questions asking: "Please list all of the advantages you perceive of preregistration", "Please list all of the disadvantages", and "Do you see any barriers to preregistration?". Participants will also be asked to confirm that they have/have not preregistered their dissertation, to reliably establish allocation of participants to each condition.

Perceptions of supervisory support. Finally, due to the literature that suggests that perceived supervisor support affects students' experiences of their dissertation research (Roberts & Seaman, 2018), to assess students' perceptions of their supervisory support at Time 2, we will use a 13-item measure of perceptions of supervisor support. This scale includes items such as "*I am satisfied with the support I have received from my supervisor*" and "*My supervisor was knowledgeable about research design/process as related to my project*.". These are measured on a 1 (strongly disagree) to 5 (strongly agree). These will be computed into one overall score of supervisory support and used as a covariate in further analyses.

Risk and Mitigations

We acknowledge certain risks associated with our study and aim to mitigate these with the following measures. The first risk is participant attrition from Time 1 to Time 2, leading to incomplete data across measures. We aim to mitigate this by accounting for average attrition rates in our planned sample as per other longitudinal studies conducted on Prolific (7%-24%; Palan & Schitter, 2018) and utilising a varied recruitment approach. Similarly, recruitment of the preregistration group requires a level of buy-in from institutions that embed a preregistration model into their undergraduate dissertation process. Members of the research

Deleted: in order

Commented [KM6]: Are there any incentives that can be offered to motivate participation at Time 2?

team have contacts with these institutions listed in Table 1, which should mitigate barriers to student access in the preregistration group.

We have also factored in discrepancies in definitions of preregistration practices, by providing all students with a student-friendly, accessible definition of preregistration from the literature (Lindsay et al., 2016). This should mean that students are able to readily identify whether they engaged in this specific process, above and beyond other processes within the dissertation timeline (e.g., discussing a protocol with their supervisor or writing an ethics application). By asking students to confirm at Time 2 that they have preregistered their study, this should also alleviate any problems with students erroneously being allocated to the wrong condition at Time 1.

Deleted: or not

Finally, our study may have confounds that we aim to reduce. For example, it is likely that institutions that actively embed preregistration into the dissertation process may also teach Open Science practices more generally within their curriculum, which may be a confound when evaluating the effectiveness of study preregistration. This will first be checked by establishing whether there are differences in students' Open Science attitudes and knowledge at Time 1. Secondly, we mitigate this by investigating the *interaction* between Group and Time on all of our outcome variables. Specifically, we expect that despite any differences between groups at Time 1, there will be a significant interaction indicating that the preregistration process has an additive effect on students' attitudes, behaviours, and perceptions of open science. Finally, we will avoid missing data adversely impacting our statistical power by using a 'forced entry' option on Qualtrics, so participants are unable to progress in the survey without completing all items in full.

Analysis Strategy

A total of X participants were excluded from analyses due to failing the attention check. Baseline characteristics of perceived supervisory support and prior statistics attainment did not

significantly differ/significantly differed between the preregistration and control group (see Table 2).

Table 2.

Baseline characteristics between the preregistration and control groups (mean and standard deviation).

	Preregistration	Control
Perceptions of Supervisor		
Prior statistics attainment		

A series of 2 (Group: preregistration vs control) x 2 (Time: time 1 vs. time 2) mixed ANOVAs will be conducted on attitudes towards statistics (SATS-28), attitudes towards QRPs, and understanding of Open Science. Any baseline differences between groups on perceptions of supervisor and prior statistics attainment will be entered as a covariate in these analyses (ANCOVA). Bonferroni corrections will be applied to elucidate pairwise comparisons, with statistical significance denoted at $p < .05$. Bayes factors will be calculated for all analyses to evaluate strength of evidence (Dienes, 2011). In line with recommendations for early research (Schönbrodt et al., 2017), we will consider $BF_{10} > 6$ as evidence for the alternative hypothesis. Bayes factors will also be used to evaluate any null results with $BF_{10} < 0.17$ considered as evidence for the null hypotheses. There is no previous literature to guide an informed prior, and thus Bayesian analyses will be computed using the default JZS prior (r = 0.707; Rouder et al., 2009) in JASP (JASP Team, 2020). The JZS prior is a noninformative default and objective prior designed to minimise assumptions about the expected effect size.

As an exploratory analysis, we will also conduct a between-participants ANOVA on Time 1 responses to the capability, opportunity and motivation (COM-B) questionnaire, to assess enablers and barriers to preregistration.

Confirmatory Analyses

Deleted:

Table 3. Need a table caption here

Research questions	Hypotheses	Sampling plan	Analysis plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
Is preregistration a useful pedagogic practice to improve students' understanding of research methods and statistics in the undergraduate dissertation?	We generally predict that attitudes to statistics will improve over time as a result of engaging with the third-year dissertation process itself, but that preregistration will have an additive effect on this. Students in the preregistration group will show a marked improvement compared to those in the control.	Two-hundred and forty final-year undergraduate Psychology students will be initially recruited with approximately 20% attrition expected at Time 2 based on prior research sampling from online platforms (Palan & Schitter, 2018). The final planned sample	2 (Group: preregistration vs control) x 2 (Time: time 1 vs. time 2) mixed ANOVA with attitudes to statistics as the dependent variable.	Simulation based power analyses conducted using the superpower shiny package (Lakens & Caldwell, 2021) with 10,000 simulations indicate that this sample size will have 80% statistical power to detect an effect size of *np2* = .04 for the two-way interaction between Group and Time, and 80% power to detect small-moderate effects of *d* = .40 for the focal pairwise comparison	This could find that preregistration *does* impact students' statistics attitudes, as we predict, or it could suggest that preregistration does not add benefits above and beyond differences that occur due to time (from time point 1 to time point 2). No main effect of time would suggest that students do not change in their attitudes towards statistics as they progress through their academic studies in final year.	Theoretically, the theory that preregistration confers a tangible, pedagogical benefit to students in their dissertation process could be unsupported by all of our proposed analyses. This may call into question various situational, contextual, and personal factors that impact how preregistration may

Kelsey McCune
This table is nice for its clear summary of the background and methods stated above, but it is cumbersome and difficult to read.
Because the Sampling plan, Rational and Theory columns are described in-depth above and apply over multiple Research Questions, I suggest condensing the table so that it contains only the Research Questions, Hypotheses, Analysis Plan, and Interpretation.

Kelsey McCune
I think this column does not need to repeat the predicted interpretation of results and should only include the interpretation given alternative results. Furthermore, although the column title indicates there will be interpretation of unexpected results, I think more detail could be added. For example, what are some reasons that preregistration might fail to improve awareness of QRPs?

Kelsey McCune
Deleted: s

Does the process of preregistration enhance awareness and endorsement of questionable research practices (QRPs)?	We predict that preregistration will reduce endorsement of QRPs as 'sensible' for the preregistration compared to the control group.	size is therefore 200 participants. See design and participants for power analysis in more detail.	2 (Group: preregistration vs control) x 2 (Time: time 1 vs. time 2) mixed ANOVA with endorsement of QRPs as the dependent variable.	between preregistration vs. control at Time 2 (Code/Output: https://osf.io/y9vz7/). We will also run a sensitivity analysis to compare our achieved sample size with planned sample size (see Participants and design section for further details).	Similarly, this analysis tests whether a preregistration process improves students' awareness of QRPs; therefore, this analysis could find that preregistration *does* positively impact students' awareness of QRPs, as we predict, or it could suggest that preregistration does not add benefits above and beyond differences that occur due to time (from time point 1 to time point 2).	be useful to students in this context.
Does the process of preregistration improve understanding of Open Science practices?	We predict that preregistration will improve understanding of Open Science practices and terminology compared to the control group.		2 (Group: reregistration vs control) x 2 (Time: time 1 vs. time 2) mixed ANOVA with awareness of Open Science		As above, this analysis allows us to test whether preregistration improves students' understanding of Open Science practices. Similar to the above, a significant main effect of Group would indicate that	

			practices as the dependent variable.		preregistration does or does not impact students' Open Science understanding, independent from time effects. Interactions of the ANOVA could find that preregistration *does* positively impact students understanding of Open Science , as we predict, or it could suggest that preregistration does not add benefits above and beyond differences that occur due to time (from time point 1 to time point 2).	
Do students recognise the benefits of the preregistration process in their undergraduate dissertation and are there any	This research question is exploratory. We will first explore whether preregistration is associated with	This research question is exploratory and the same sample detailed above will be used to	A between-participants (Group: preregistration vs control) ANOVA with	This research question is exploratory. Qualitative research typically does not share concerns of generalisability with	This set of exploratory analyses allows us to test whether students have the sufficient capability, opportunity, and motivation to complete	

barriers/challenges to its implementation?	*Capability, Opportunity, and Motivation* (COM-B) *for preregistration* by comparing the preregistration and control groups. We will then conduct qualitative content analysis on participant's free-text responses at Time 2.	address this question.	COM-B scores as the dependent variable. Qualitative analysis using qualitative content analysis for free-text responses.	quantitative research, so our planned sample size for this study will be sufficient for our qualitative research question, given the epistemological underpinnings of this approach.	preregistration. Qualitative analyses will shine light into whether students recognise any barriers or challenges, in order to provide more nuance to the quantitative analysis.	

Commented [KM9]: Although this research question is exploratory, I think the results will be important for understanding how or whether to modify specific components of the preregistration process to increase the ease and motivation of early-career researchers in engaging with open science. Therefore, here and in the hypotheses column I would like to see more details on what specifically you expect from the analysis comparing the preregistration and control group of students.

Attitudes Toward Statistics

We predict that there will be a main effect of time, in that over time students' perceptions of statistics will improve (i.e. their scores on this scale will go down) in both groups (Table 3). We also predict that there will be a two-way interaction between Group and Time with the preregistration condition exerting an *additive* effect on this to show more marked improvement in statistics attitudes. In line/contrary to hypotheses, there was a [significant/no significant] main effect of Group and Time, and a [significant/non-significant] two-way interaction between Group and Time on *statistics affect/cognitive competency/value/difficulty*. Pairwise comparisons indicate that...

Attitudes towards QRPs

Since the other research questions include summarized predictions, it is important to include that here as well. In line/contrary to hypotheses, there was a [significant/no significant] main effect of Group and Time, and a [significant/non-significant] two-way interaction between Group and Time on attitudes towards QRPs. Pairwise comparisons indicate that...

Deleted: ↵

Understanding of Open Science

Unlike with statistics attitudes, we do not expect that there will be a main effect of time, that is, participants who have not preregistered will not show an improvement in Open Science understanding. We predict an preregistration condition * time interaction, whereby participants in the preregistration group will improve their understanding from Time 1 to Time 2. In line/contrary to hypotheses, there was a [significant/no significant] main effect of Group and Time, and a [significant/non-significant] two-way interaction between Group and Time on understanding of Open Science. Pairwise comparisons indicate that…

Deleted: that

Exploratory Analyses

COM-B

A between-participants ANOVA indicated a [significant/no significant] main effect of Group on *Capability, Opportunity, and Motivation for preregistration.* Those in the preregistration group showed... compared to the control.

Qualitative analysis

Students' responses to the open-ended questions at Time 2 will be analysed using thematic analysis (Braun & Clarke, 2006). This will involve one author reading and coding the free-text responses for their content before discussing potential codes with the rest of the authorship team. One author will then code the full set of verbatim textual responses and generate themes and subthemes on the data.

Discussion

Here, we will discuss our results as they relate to the utility and effectiveness of embedding preregistration into the undergraduate dissertation process.

Limitations

There are certain limitations that we recognise prior to data collection. For example, students and supervisors who develop a detailed, rigorous preregistration and engage in the process more with their supervisor might report greater benefits compared to those who develop a poor quality, less detailed preregistration. However, it is beyond the scope of this research to assess each preregistration for quality and rigour. Therefore, future work, depending on our findings, may wish to establish the extent to which preregistration quality impacts on the core outcomes of interest in this work.

References

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., ... & Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *27*(2), 108-119.

Ayebo, A., Bright, J., & Ballam, C. (2019). Examining the factor structure of the survey of attitudes towards statistics among undergraduate health science students. *International Electronic Journal of Mathematics Education*, *15*(1), em0560.

Banks, G. C., Rogelberg, S. G., Woznyj, H. M., Landis, R. S., & Rupp, D. E. (2016). Editorial: evidence on questionable research practices: The good, the bad, and the ugly. *Journal of Business and Psychology,* 31(3), 323–338. https://doi.org/10.1007/s10869- 016-9456-7

Blincoe, S., & Buchert, S. (2020). Research preregistration as a teaching and learning tool in undergraduate psychology courses. *Psychology Learning & Teaching*, *19*(1), 107-115.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Button, K. S., Chambers, C. D., Lawrence, N., & Munafò, M. R. (2020). Grassroots training for reproducible science: a consortium-based approach to the empirical dissertation. *Psychology Learning & Teaching*, *19*(1), 77-90.

Button, K. S., Lawrence, N. S., Chambers, C. D., & Munafò, M. R. (2016). Instilling scientific rigour at the grassroots. *The Psychologist*, *29*(3), 158-159.

Button, K. (2018). Reboot undergraduate courses for reproducibility. *Nature*, *561*(7723), 287-288.

Chopik, W. J., Bremner, R. H., Defever, A. M., & Keller, V. N. (2018). How (and whether) to teach undergraduates about the replication crisis in psychological science. *Teaching of Psychology*, *45*(2), 158-163.

Creaven, A. M., Button, K., Woods, H., & Nordmann, E. (2021). Maximising the educational and research value of the undergraduate dissertation in psychology. *PsyArXiv*

Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J. C., ... & Schulte-Mecklenbeck, M. (2019). Seven Easy Steps to Open Science. *Zeitschrift für Psychologie.*

Day, J., & Bobeva, M. (2007). Applying performance management principles to a learning environment for undergraduate dissertations: A case study. *International Journal for Quality and Standards*, *1*(1), 217-239.

de Leeuw, J. R., Andrews, J., Livingston, K., Franke, M., Hartshorne, J., Hawkins, R., & Wagge, J. (2019). Using replication studies to teach research methods in cognitive science. *Perspectives on Psychological Science*, *7*(6), 600-604.

Dauphinee, T. L., Schau, C., & Stevens, J. J. (1997). Survey of Attitudes Toward Statistics: Factor structure and factorial invariance for women and men. *Structural Equation Modeling: a Multidisciplinary Journal*, *4*(2), 129-141.

Detweiler-Bedell, B., & Detweiler-Bedell, J. B. (2019). Undergraduate research teams that build bridges, produce publishable research, and strengthen grant proposals. *Frontiers in Psychology*, *10*, 133.

Devezer, B., Navarro, D. J., Vandekerckhove, J., & Buzbas, E. O. (2020). The case for formal methodology in scientific reform. *BiorXiv*.

Devonport, T. J., & Lane, A. M. (2006). Cognitive appraisal of dissertation stress among undergraduate students. *The Psychological Record*, *56*(2), 259-266.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on?. *Perspectives on Psychological Science*, 6(3), 274-290.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502-1505.

Frank, M. F., & Saxe, R. (2012). Teaching replication. *Perspectives on Psychological Science,* 7, 600–604. https://doi.org/10.1177/1745691612460686

Frith, U. (2020). Fast lane to slow science. *Trends in cognitive sciences*, *24*(1), 1-2.

Greenbank, P., Penketh, C., Schofield, M., & Turjansky, T. (2008). The undergraduate dissertation: 'Most likely you go your way and I'll go mine'. *The International Journal for Quality and Standards*, *3*(22), 1-24.

Grimes, D. R., Bauch, C. T., & Ioannidis, J. P. (2018). Modelling science trustworthiness under publish or perish pressure. *Royal Society Open Science*, *5*(1), 171511.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biol*, *13*(3), e1002106.

Janz, N. (2016). Bringing the gold standard into the classroom: replication in university teaching. *International Studies Perspectives*, *17*(4), 392-407.

JASP Team. (2020). JASP (Version 0.14.1) [Computer software].

Jekel, M., Fiedler, S., Allstadt Torras, R., Mischkowski, D., Dorrough, A. R., & Glöckner, A. (2020). How to Teach Open Science Principles in the Undergraduate Curriculum—The Hagen Cumulative Science Project. *Psychology Learning & Teaching*, *19*(1), 91-106

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*(5), 524-532.

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*(3), 196-217.

Keyworth, C., Epton, T., Goldthorpe, J., Calam, R., & Armitage, C. J. (2020). Acceptability, reliability, and validity of a brief measure of capabilities, opportunities, and

motivations (“COM-B”). *British Journal of Health Psychology,* 25, 474-501. https://doi.org/10.1111/bjhp.12417

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L. S., Kennett, C., Slowik, A., Sonnleitner, C., Hess-Holden, C., Errington, T., Fielder, S. & Nosek, B. A. (2016). Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency. *PLoS biology*, *14*(5), e1002456.

Klein, O., Hardwicke, T. E., Aust, F., Breuer, J., Danielsson, H., Mohr, A. H., ... & Vazire, S. (2018). A practical guide for transparency in psychological science. *Collabra: Psychology*, *4*(1).

Krishna, A., & Peter, S. M. (2018). Questionable research practices in student final theses – Prevalence, attitudes, and the role of the supervisor’s perceived attitudes. *PLoS One,* 13, e0203470. https://doi.org/10.1371/journal.pone.0203470

Kvetnaya, T., Frank, M., Brachem, J., Hill, M., Falk Florentin Schramm, L., & Eiberger, A. (2019). Questionable research practices and open science in undergraduate empirical projects: Results from a nationwide survey amongst German psychology students. *Presentation at the Proceedings of Open Practices in Education conference.* http://www.dipfdocs.de/volltexte/2020/19034/pdf/Heck_2019_OPINE_Proceedings_A.pdf#page=42

Lakens, D. (2021). Sample size justifications. *PsyArXiv* https://psyarxiv.com/9d3yf/

Lakens, D., & Caldwell, A. (2021). Simulation-based power analysis for factorial Analysis of Variance designs. *Advances in Methods & Practices in Psychological Science*, 4, 1-14. https://doi.org/10.1177/2515245920951503

LeBel, E. P., & Peters, K. R. (2011). Fearing the future of empirical psychology: Bem's (2011) evidence of psi as a case study of deficiencies in modal research practice. *Review of General Psychology*, *15*(4), 371-379. https://doi.org/10.1037/a0025172

LeBel, E. P., Borsboom, D., Giner-Sorolla, R., Hasselman, F., Peters, K. R., Ratliff, K. A., & Smith, C. T. (2013). PsychDisclosure. org: Grassroots support for reforming reporting standards in psychology. *Perspectives on Psychological Science*, *8*(4), 424-432.

Lindsay, D. S., Simons, D. J., & Lilienfeld, S. O. (2016). Research preregistration 101. *APS observer*, *29*(10).

Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1), 1-12.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Du Sert, N. P., ... & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1-9. https://doi.org/10.1038/s41562-016-0021

Nicholas, D., Watkinson, A., Boukacem-Zeghmouri, C., Rodríguez-Bravo, B., Xu, J., Abrizah, A., ... & Herman, E. (2017). Early career researchers: Scholarly behaviour and the prospect of change. *Learned Publishing*, *30*(2), 157-166. https://doi.org/10.1002/leap.1098

Norris, E. & O'Connor, D. B. (2019) Science as behaviour: Using a behaviour change approach to increase uptake of open science, *Psychology & Health*, 34(12), 1397-1406, https://doi.org/10.1080/08870446.2019.1679373

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600-2606.

Nuzzo, R. (2015). How scientists fool themselves–and how they can stop. *Nature News*, *526*(7572), 182.

O'Boyle Jr, E. H., Banks, G. C., & Gonzalez-Mulé, E. (2017). The chrysalis effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management*, *43*(2), 376-399.

Olson, J., Mosen, J., Voracek, M., & KIrchler, E. (2019). Research practices and statistical reporting quality in 250 economic psychology master's theses: A meta-research investigation. Royal Society Open Science, 6, 190738. httpL//dx.doi.org/10.1098/rsos.190738

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics Anxiety: Nature, etiology, antecedents, effects, and treatments--a comprehensive review of the literature. *Teaching in Higher Education*, *8*(2), 195-209.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).

Orben, A. (2019). A journal club to fix science. *Nature*, *573*(7775), 465-466.

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22-27.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence?. *Perspectives on Psychological Science*, *7*(6), 528-530.

Pownall, M. (2020a). Pre-registration in the undergraduate dissertation: A critical discussion. *Psychology Teaching Review*, *26*(1), 71-76.

Pownall, M. (2020b) Early career researchers in open science: Vanguard or cannon fodder? *The Psychologist* .https://thepsychologist.bps.org.uk/early-career-researchers-open-science-vanguard-or-cannon-fodder

Roberts, L. D., & Seaman, K. (2018). Good undergraduate dissertation supervision: perspectives of supervisors and dissertation coordinators. *International Journal for Academic Development*, *23*(1), 28-40.

Rouder, J. Speckman P. Sun D. Morey R. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 255-237. https://doi.org/10.3758/pbr.16.2.225

Sacco, D. F., & Brown, M. (2019). Assessing the efficacy of a training intervention to reduce acceptance of questionable research practices in psychology graduate students. *Journal of Empirical Research on Human Research Ethics,* 14, 209-218. https://doi.org/10.1177/1556264619840525

Sarafoglou, A., Hoogeveen, S., Matzke, D., & Wagenmakers, E-J. (2020). Teaching good research practices: Protocol of a research master course. *Psychology, Learning & Teaching*, 19, 46-59. https://doi.org/10.1177/1475725719858807

Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, 22(2), 322.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359-1366.

Sorokowski, P., Groyecka, A., Błaszczyński, K., Frąckowiak, T., & Kobylarek, A. (2019). Registered reports as a method to increase credibility of science–experimental study among psychological students. *Journal of Education Culture and Society*, *10*(2), 67-75.

Stürmer, S., Oeberst, A., Trötschel, R., & Decker, O. (2017). Early-career researchers' perceptions of the prevalence of questionable research practices, potential causes, and

open science. *Social Psychology*. 48(6), 365-371. https://doi.org/10.1027/1864-9335/a000324

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. (2016). The academic, economic and societal impacts of Open Access: an evidence-based review. *F1000Research*, *5*.

van't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology*, *67*, 2-12.

Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, *7*(6), 632-638.

Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019). Publishing research with undergraduate students via replication work: The collaborative replications and education project. *Frontiers in Psychology*, *10*, 247.

Zečević, K., Houghton, C., Noone, C., Lee, H., Matvienko-Sikar, K., & Toomey, E. (2020). Exploring factors that influence the practice of Open Science by early career health researchers: a mixed methods study. *HRB Open Research*, *3*(56), 56.