An empirical investigation into the social, psychological, and structural correlates of researcher misbehaviour

Robin Brooker

A thesis submitted for the degree of PhD in Sociology

Department of Sociology and Criminology

University of Essex

Date of submission for examination (September 2025)

Date of Corrections (May 2026)

Acknowledgements

I would like to extend my deepest gratitude to my supervisors, Professor Nick Allum and Dr Laurie James-Hawkins. Their guidance, encouragement, and wisdom have been invaluable throughout the course of this PhD, and I have learned so much from them about what it truly means to be a good academic. Nick, thank you for introducing me to so many inspiring scholars and for providing countless opportunities to learn, network, and grow. Laurie, I am especially grateful for your patient efforts to teach me the importance of saying "no" – it took some time, but I think I am finally getting there.

Thank you to my partner, Cass, for your unwavering support and for being my greatest champion in every aspect of life. I am endlessly grateful for all that you have done – especially for serving as my ever-patient "proofreader-in-chief."

Thank you to my parents for instilling in me a love of learning and sparking my intellectual curiosity. I am especially grateful to my dad, who sadly was not here to see me begin or complete this PhD, but whose pride I know would have been immense.

Summary of Thesis

This thesis examines the systemic, institutional, and individual-level correlates of questionable research practices (QRPs) across three empirical papers. Collectively, the papers contribute to the development of a comprehensive explanatory framework for understanding researcher misbehaviour across disciplinary and national contexts.

Paper A draws on secondary analysis of the *International Research Integrity Survey* (Allum et al., 2023) to assess the relative association with individual- and institutional-level factors on self-reported QRP engagement. Employing a multilevel modelling approach, this study identifies key predictors of research practices at the measurement, individual and institutional levels. Paper A has been published in *Research Integrity and Peer Review* (Brooker & Allum, 2024).

Paper B builds on these findings by investigating one potential mechanism underlying self-reported QRP engagement: researchers' perceptions of their peers' behaviours within their field. Using secondary data from a large-scale international survey conducted as part of the *Practices, Perceptions, and Patterns of Research Integrity* (PRINT) project, Paper B employs Bayesian multilevel modelling to test whether perceptions of peer behaviour predict individual self-reported engagement in QRPs.

Paper C extends this line of inquiry by examining additional individual- and system-level correlates of QRPs. Specifically, it explores the roles of researchers' values, motivations, normative commitments, epistemic orientations, personality traits, and political orientations. Unlike the preceding studies, the data analysed in Paper C were collected specifically for this thesis and comprise survey responses from researchers working across a range of disciplines and countries.

Table of Contents

Chapter 1

Introduction

Introduction

Replication *is* the cornerstone of science. When a scientific finding is real, teams of scientists *should* be able to run the same study and reproduce the same results. If this attempt at replication fails, and the same results are not reproduced, the scientific record should be corrected. In other words, self-correction represents the primary goal of science – whereby errors are weeded out over time – and replication represents the primary means of achieving self-correction. However, since 2010, many exciting, newsworthy findings have not held up when subjected to empirical scrutiny. Take 'power posing', for example. This practice, of adopting an expansive, open body posture, was widely promoted as a means of boosting confidence, feelings of power and risk-taking behaviour. However, large-scale replication attempts failed to reproduce these effects. Cases like this have raised an important question: if replication is the cornerstone of science, why do some findings not stand up to repeated empirical scrutiny?

This thesis addresses one critical dimension of the answer: questionable research practices (QRPs). QRPs are situated in the ambiguous area between rigorous scientific practice and outright misconduct. They encompass behaviours such as hypothesising after results are known, continuing data collection until statistical significance is reached, or failing to report non-significant findings. Such practices are frequently employed without intent to deceive and in some cases may be defensible on ethical or empirical grounds. However, researchers are embedded in systems that privilege novelty, innovation, and positive outcomes, they work at institutions with specific cultures, policies and rules, and they have their own motivations, values, traits and beliefs about how science ought to be done. The central argument of this thesis is that QRPs emerge as the product of interactions between these levels: systemic incentives, local research climates, and individual dispositions.

On this basis, this thesis asks three questions. Firstly, to what extent does self-reported QRP engagement vary from place to place – between countries and research institutions – and which institutional features matter? (Paper A) Secondly, do researchers' perceptions of what others are doing correlate their own behaviour – are they more likely to cut corners and cheat because their peers do? (Paper B) Thirdly, how do individual factors (like values, personality traits, and motivations) stack up against systemic pressures (publication and funding pressure) in predicting self-reported QRP engagement? (Paper C).

The downstream consequences of engaging in QRPs are profound. Findings built on fragile foundations *can* enter the scientific record, resources of time, money, and intellectual effort are misdirected toward effects that do not exist, policy responses are shaped by misleading evidence, and public trust in science is undermined. Other QRPs, such as gift authorship or excluding important information from a scientific manuscript, can have other effects, such as undermining accountability, and potentially impacting the perceived legitimacy of science. At the very least, such practices are disingenuous. Understanding the correlates of QRPs is therefore essential for identifying appropriate points of intervention. If these practices are primarily a product of systemic incentives, reform in how scientific contributions are rewarded and evaluated is required. If they cluster within particular institutions, changes to research culture, governance, and institutional policy are necessary. If, however, individual motivations, values, and normative commitments are decisive, then interventions must target researchers themselves. This thesis advances a framework of the correlates of self-reported QRP engagement in order to identify the most efficacious leverage points for enhancing the reliability, trustworthiness, and societal impact of science.

Background

A plethora of scientific conclusions have now attained an established status within the accepted canon of scientific knowledge. For example, there is little room for doubt that cigarette smoking is a powerful, dose-dependent cause of lung cancer (Lee et al., 2012; Gandini et al., 2008; Feiling et al., 2023; O'Keeffe et al., 2018; Zeng et al., 2019). Moreover, we know from cognitive psychology that people take longer and are more error-prone when naming the ink-colour of an incongruent colour-word (e.g., the word 'red' printed in blue; Macleod, 1991) – a phenomenon known as the Stroop effect (Stroop, 1935). We also know from the field of Economics that for each additional year of school, there is average causal gain of approximately 9% in annual earnings (Psacharopoulos & Patrinos, 2018; Horie & Iwasaki, 2022; Montenegro & Patrinos, 2014). These established 'facts' – that is, empirical observations that have been consistently confirmed and widely accepted as true – have informed effective interventions, regulation, and policies, that have successfully and tangibly impacted society in their respective domains. For example, the well-established link between cigarette smoking and lung cancer lead to restrictions on advertisement and marketing, increased taxation, widespread smoking bans in certain indoor spaces, numerous public health campaigns, and litigation against tobacco companies. As a result, there was measurable decline in smoking – in the UK, 20.2% of the UK population smoked in 2011, compared to 11.9% in 2023, a decline of 8.3 percentage points over a ten-year period (Office for National Statistics, 2024).

What unites these diverse phenomena is their empirical validation across multiple, independent studies that have reproduced the original effects. Through such successive replications, findings acquire the robustness needed to enter the canon of scientific knowledge (Schmidt, 2009). Yet concerns have grown in recent years that this self-correcting process is faltering, with many findings failing to replicate. This so-called "replication crisis" (Baker, 2016; Pashler & Harris, 2012; cf. Rawat, 2022) has been most visible in psychology and medicine, where high-profile replication initiatives have produced sobering results. This awareness has in part been driven by several high-profile replication efforts that did not yield statistically significant results or effect sizes consistent with the original findings (Open

Science Collaboration, 2015; Camerer et al., 2018). For example, an early, large-scale reproducibility attempt in psychology undertaken by Brian Nosek and colleagues, attempted to replicate one-hundred studies published in the *Journal of Personality and Social Psychology*, *Journal of Experimental Psychology: Learning, Memory & Cognition*, and *Psychological Science*. Of the ninety-seven original articles that found statistically significant effects, only 36% of these effects could be replicated (Open Science Collaboration, 2015). Similarly, Camerer et al. (2018) attempted to replicate the findings from twenty-one social and behavioural science papers published in *Nature* and *Science* between 2010 and 2015. Statistically significant effects in the same direction of the original effect were only found in 62% of the studies.

These replication failures have prompted intense debate about their underlying causes. The answer is multi-factorial, with a combination of statistical, methodological, and sociological reasons. Some phenomena may exist only within specific cultural, temporal, or procedural contexts (Klein et al., 2018), meaning attempts to replicate are destined to fail because they are context-specific. For example, some findings may result from hidden moderators, with effect-size heterogeneity only revealed in diversified replications. Likewise, confounders may remain unknown, unmeasured, or measured poorly, meaning that original estimates are biased and spurious. If an effect is identified among a sample of undergraduate students in the United Kingdom in 2005, a replication among a sample of online adults in China in 2024 may fail to find the same effect, on the basis that the effect is context or temporally specific. Researchers rarely know beforehand whether these contextual factors matter, with a default to *ceteris paribus* assumptions, whereby researchers only engage in boundary-testing once the base effect appears secure. This represents the ‘normal’ process of scientific enquiry, whereby researchers start by exploring an empirical claim, before testing its robustness and engaging in boundary-mapping through replication. On this basis, failures to replicate may sometimes represent a benign and integral part of the normal process of science and are not a cause for alarm.

That being said, there is an emerging consensus within the meta-scientific community that many of these replication failures can, at least in part, be attributed to causes that are more insidious. At the sharper end lies overt scientific misconduct, defined as fabrication, falsification, and plagiarism (FFP; John et al., 2012). While serious, misconduct appears comparatively rare, with prevalence estimates for falsification and fabrication ranging from 1.9% to 4.3% depending on the study (Fanelli, 2009; Gopalakrishna et al., 2022). More robust field-specific evidence comes from forensic analyses, such as image-screening studies in biomedical research, which identify problematic duplications in 3.8% of papers (Bik et al., 2016; Bucci, 2018). Retractions provide further evidence, although they remain infrequent (0.05–0.2% of published papers in 2023; Van Noorden, 2023), and many result from errors rather than misconduct. Even so, misconduct is implicated in around 60% of biomedical retractions (Hwang et al., 2023).

However, there is widespread consensus that the replication crisis has less to do with overt cases of scientific misconduct and more to do with questionable research practices (QRPs; John et al., 2012; Simmons et al., 2011; Wicherts et al., 2016).[1] Beyond replicability,

[1] It is important to note that the relationship between QRPs and replication is not straightforward. Some research suggests that QRPs do not necessarily have large effects on replication outcomes (Ulrich & Miller, 2020), that their effects may be limited to particular research contexts (Rubin, 2023; Rubin, 2025), and that, in some cases, they may even increase apparent replication rates (Francis, 2012). Commonly discussed QRPs, such as p-hacking, HARKING, and publication bias, may therefore be less effective in explaining the replication crisis than is often assumed (Rubin, 2025), and may not always be as problematic as previously suggested (Leung, 2011; Oberauer & Lewandowsky, 2019; Mohseni, 2020; Rubin, 2022; Vancouver, 2018; Adda et al., 2020; Erasmus, 2024; Gupta & Bosco, 2023; Rooprai et al., 2023). Moreover, the extent to which low replication rates should themselves be regarded as a problem has also been contested (Rubin, 2023; Rubin, 2025; Fanelli, 2018; Lewandowsky & Oberauer, 2020).

The author recognises that *not* all QRPs are unethical or harmful in every instance. Given their ethical ambiguity, and the fact that some may be justifiable in particular contexts, QRPs have other potential downstream effects, such as wastage of resources, growth of false research literatures, erosion of trust, reduced credibility of legitimate findings and slower self-correction.

QRPs are without straightforward definition but include practices that exist in the ethically ambiguous area between sound scientific practice (characterised by methodologically robust, ethical, transparent, and open inquiry) and scientific misconduct. They are numerous and involve the exploitation of the flexibility that exists at various stages of the research process, from research, through peer review, to publication. They include practices such as p-value hacking (or, 'p-hacking'), which is a group of practices that involve running many different statistical tests and analytical specifications (e.g., stopping data collection early, including/removing covariates, running multiple modelling choices) until statistical significance is obtained (Head et al., 2015; Norman, 2014), which increase the likelihood of a Type-I error rates (Simmons et al., 2011) and exaggerate effect sizes. Relatedly, Hypothesising After Results are Known ('HARKing'; Kerr, 1998), involves passing off post-hoc findings as if they are an a-priori theory. On the one hand, this is disingenuous, but on the other hand, hypotheses are retrofitted to data, so they cannot be proven wrong, undermining the scientific goal of Popperian falsification. Other examples from quantitative research include under-powered studies, lack of blinding or randomisation in experiments and not correcting for multiple statistical tests (e.g., by using Bonferroni corrections).

---

certain QRPs may sometimes be appropriate. However, even where QRPs have limited effects on replication rates, they may still have important downstream consequences, including the production of false research literatures, wasted resources, distorted theory development, poor educational signals, and reduced credibility of legitimate findings. On this basis, the present discussion treats QRPs not merely as contributors to replication failure, but as practices that may undermine the integrity and reliability of research more broadly.

QRPs are not limited to quantitative research either and extend beyond the data collection and analysis stages, with other QRPs including misleading grant or ethics applications (Bouter et al., 2019), poor record keeping, concealing null results in publications, over-claiming, gifting authorship, excluding authors where authorship is warranted, citing one's own work when it is irrelevant, or ignoring legitimate requests for publication correction. Emerging QRPs include not making code and data available in accordance with open science mandates and policies, and using – or, not indicating use of – generative language AI in research design, analysis, manuscript preparation and peer review. While these practices, when in isolation, are less impactful than misconduct, the concern arises given the frequency at which researchers engage in them (John et al., 2012; cf. Fiedler & Schwarz, 2016). One of the earliest studies investigating the prevalence of QRPs was a meta-analytical study conducted by Martinson et al. (2005). In their survey study, they found that approximately 33% of researchers admit to having engaged in a QRP or misconduct within the three years prior to completing survey. A subsequent meta-analysis found that 33.7% of researchers admit to having engaged in QRPs (Fanelli, 2009). Recognising that such self-report estimates likely underestimate the true extent of self-reported engagement – given response bias when admitting to sensitive and unethical behaviours – John et al. (2012) used incentives for truth telling and an anonymous elicitation survey format to assess QRP prevalence estimates amongst psychologists. They found prevalence estimates up to 100% for some QRPs, including p-hacking and selectively-reporting. Conversely, a more recent meta-analytical study by Xie et al. (2021) found more modest prevalence estimates of 12.5%. However, given that self-reported engagement is non-zero, the cumulative impact of these practices over time should not be understated.[2]

[2] The objective of this thesis is to investigate the relationships and associations between variables, rather than to provide a precise or definitive estimate of the population-level prevalence of QRPs. A substantial body of literature has attempted to quantify the exact baseline rates of QRP engagement (see, Martinson et al., 2005; Fiedler & Schwarz, 2016; Fanelli, 2009; John et al., 2012; Xie et al., 2021), establishing an absolute percentage is highly fraught. As discussed in the discussion of this thesis, methodologically, survey data in this domain is inherently prone to social desirability bias, differential non-response, and

The exploitation of the degrees of freedom in the research process inflate the probability of Type-I errors (that is, 'false positives'). In a demonstration using computer simulations of experimental data in psychology, it was found that merely adjusting four degrees of freedom in the analytical process – e.g., analysing multiple dependent variables, increasing the number of observations – statistical significance can be found in up to 61% of cases where no effect actually exists, meaning a twelve-fold inflation in the Type-I error rate (Simmons et al., 2011). In a landmark paper, John Ioannidis presented a positive-predictive-value model showing that when this bias is layered on to weak priors and low power, the post-study probability that a statistically significant finding is in fact true can fall below the 50% mark (Ioannidis, 2005). This provided basis for Ioannidis' claim that 'most scientific findings are false'. When a follow-up study is run to assess the original effect under a strict, pre-planned analytical strategy that does not allow for the exploitation of the flexibility in research process, the original effect almost always reduces in magnitude or disappears. However, erroneous effects and invalid theories can persevere for considerable time, as biases in the publication system that favour statistically significant, novel, and exciting findings may mean the knowledge-correcting findings are filled away and not published (i.e., 'the file drawer' problem; Rosenthal, 1979; Sterling et al., 1995). The visible scientific literature is pre-filtered for success, meaning effect sizes are inflated upwards and theoretical claims appear more robust than they really are.[3]

---

varying interpretive frames across different disciplines. Fundamentally, this thesis operates on the principle that from an institutional and ethical standpoint, the absolute prevalence rate is secondary to the fact that the baseline is non-zero. Whether a specific practice is reported by 10% or 50% of a given sample does not alter the underlying reality that QRPs introduce systemic unfairness, compromise scientific integrity, and reward disingenuous behaviours. The primary harm of QRPs exists independently of their exact frequency or their eventual downstream consequences on replication outcomes.

[3] Other QRPs do not impact the replicability of past results. For example, gifting authorship to an individual who has not contributed sufficient to the project dilutes accountability, credit and resources become disingenuously misallocated, and post-publication dialogue becomes difficult and less-efficient (contact authors may not have access to data or code, thereby undermining reproducibility, and requests for corrections may hit dead ends).

The epistemic fallout is severe. Spurious effects remain canon until enough contrary evidence accumulates, or until a systematic, high-powered replication study (of the kind mentioned earlier) demonstrate the invalidity of the original claims. Significant quantities of resources, including time, money and research effort are wasted carrying-out follow-up studies that chase statistical artefacts. Chalmers & Glasziou (2009) estimate that up to 85% of global biomedical-research expenditure is wasted each year as a result of flawed design, analysis, and reporting. Downstream, policies, interventions, guidelines, and treatments informed by shaky, invalid evidence become destined to fail in addressing societal issues. In turn, these practices – and the resulting failed replications – fuel scepticism in science, and lead to a decline in trust in science, thus undermining future interventions and strategies that rely on buy-in from the general public. Experimental evidence shows that failed replications measurably lower laypeople's trust in researchers and science as an institution (Hendriks et al., 2020).

This raises a critical question: what factors lead researchers to adopt practices that compromise the validity of their findings, undermine reproducibility, diminish the societal value of their work, and erode public trust in science? An emerging, cross-disciplinary literature has sought to elucidate these causal factors and mechanisms, with specific causes identified at the individual, institutional and system-level. However, this literature is heavily fragmented, and an integrative account of why researchers, across contexts, adopt QRPs is noticeably absent.

System-level Factors

The predominant paradigm explaining *why* researchers' engage in QRPs posits that current incentive structures within academia are fundamentally misaligned with the normative ideals underpinning rigorous scientific conduct (John et al., 2012; Bouter, 2015; Edwards & Roy, 2017). Academic institutions use simple metrics for assessing for hiring, promotion, permanency, and tenure (Moher et al., 2018). This, combined with a publishing bias that favours statistically significant, novel, and exciting scientific findings (Sterling, 1959; Dwan et al., 2013; Chavalarias et al., 2016), creates a high-pressure, hyper-competitive academic environment which incentivises researchers to ‘game the system’ and ‘cut corners’ to obtain positive, aesthetically adequate and, hence, more-publishable results. That is, cheating maximises their chances – and thereby their frequency – of publication (‘publish or perish’; Head et al., 2015; Simmons et al., 2011). The favouring of positive findings by academic journals is evidenced by the visible absence of null findings in the published literature (Chavalarias et al., 2016). Researchers’ behaviours are steered by what they view as important, and avoiding joblessness (or, careerlessness) is feasibly, and understandably, held in high regard by most researchers.

Certainly, there is accumulating empirical evidence indicating a direct relationship between publication pressure and self-reported engagement in QRPs (Schneider et al., 2024; Gopalakrishna et al., 2022; Maggio et al., 2019). For example, Schneider and colleagues (2024) found that across a sample of researchers in Europe and the United States of America, publication pressure was *the* strongest predictor of self-reported QRP engagement, compared to personality traits, perceived QRP acceptability, academic age, and research culture. The number of researchers who report feeling this pressure has been found to be high, with approximately 72% of biomedical professors in the Netherlands reporting that there is ‘excessive pressure to publish’ (Tijdink et al., 2013); Maggio et al., 2019). For example, Schneider and colleagues (2024) found that across a sample of researchers in Europe and the United States of America, publication pressure was *the* strongest predictor of self-reported QRP engagement, compared to personality traits, perceived QRP acceptability, academic age, and research culture. The number of researchers who report feeling this pressure has been found to be high, with approximately 72% of biomedical professors in the Netherlands reporting that there is ‘excessive pressure to publish’ (Tijdink et al., 2013)

Consequently, reform initiatives have targeted interventions on two fronts. On the one hand, the metrics used to assess for hiring, promotion, permanency and tenure by academic institutions have received robust challenge (Hicks et al., 2015; Kretser et al., 2019; Moher, 2018). Using quantitative metrics, such as publication counts, citation counts and journal impact factors as means of ranking researchers (i.e., when ranking job candidates), fuels hyper-competition (Edwards & Roy, 2017). These measures have become 'targets', and as a result, in accordance with 'Goodharts Law', cease to function as effective indicators of quality (Fire & Guestrin, 2019). According to their analysis of over 120 million scientific articles, Fire & Guestrin (2019) find that longer author lists, shorter papers, and a monumental growth in the number of publications, have resulted in author publication counts no longer being valid, nor reliable, measures of academic success. Similarly, the authors find that citation-based metrics of researcher productivity have become compromised by the number of papers, lengthy reference lists and self-citation. As a result, alternative metrics for assessing researcher impact, such as new indices (Bartneck & Kokkelmans, 2010; Post et al., 2018; Romanovsky, 2012; Wu, 2009; Waltman et al., 2011), qualitative metrics (Hicks et al., 2015; e.g., narrative CVs; Bordignon et al., 2023), and researchers' broader contributions (e.g., replication attempts, peer review and knowledge sharing; Moher et al., 2020) are increasingly being adopted as more robust measures of scientific success.

On the other hand, there have been calls for reform of the antiquated publication model, advocating for the increased publication of non-significant results, negative effects (Ware & Munafò, 2015; Nosek, Spies & Motyl, 2012) and replication studies. Relatedly, the use of registered reports – whereby, journal editors agree to publish unknown results based on research design and methodological rigour ahead of data collection – can be used to mitigate the publication bias towards statistically significant findings. A similar approach, 'result-blind peer review' – where reviewers review an article by looking at the research question and methods *only* – is a novel means of addressing the bias towards statistically significant results (Woznyj et al., 2018). Registered reports are already being used in *Cortex* and *Comprehensive Results in Social Psychology*, while results-blind peer review trials have been undertaken by *BMC Psychology*. There are additionally journals now dedicated to publishing negative and null results, including the *Journal of Negative Results in BioMedicine* and *Journal of Articles in Support of the Null Hypothesis*. Many of these reforms are recent, so it is too early to measure how they have impacted the skew towards positive results in the scientific literature

Institutional-level Factors

As discussed, evidence suggests that institutional assessments for employment, permanency, and promotion in combination with systemic biases inherent within the academic publication system, collectively exacerbate researchers' self-reported engagement in QRPs. However, institutional research cultures, coupled with policy and governance structures – and researchers' perceptions of them – additionally shape the degree to which individuals may be willing to compromise their integrity. For example, Martinson et al. (2006), based on a survey of researchers funded by the National Institutes of Health (NIH), found that when researchers believe they are being treated unfairly by their organisation ('organisational (in)justice'), they are more likely to compromise their research integrity. Similarly, where research climates are perceived more negatively (Haven et al., 2019; Crain et al., 2014; Armond & Kakuk, 2022), work relationships are poor (as a result of competition; Anderson et al., 2007a) and mentoring is inadequate (Anderson et al., 2007b), researchers are more likely to compromise their integrity (Crain et al., 2014; Armond & Kakuk, 2022).

Moreover, research institutions have increasingly implemented policies and programmes explicitly aimed at promoting research integrity. Institutional responsibilities typically encompass providing education and training in the responsible conduct of research, maintaining clear procedures for addressing allegations of scientific misconduct, and actively encouraging best practices in scientific research. Institutions additionally respond to research-related legal and regulatory compliance responsibilities, such as regulations governing research on humans and animals, human tissue, data use, environmental regulations, and financial conduct. However, due to their inherent ethical ambiguity, QRPs frequently fall outside the scope of formal institutional (and national) governance frameworks and regulation and thus tend to be addressed primarily through general guidance provided in training sessions, mentoring relationships and indirectly through promotion of positive research cultures. That being said, the open science movement has seen widespread implementation of policies and practices based on national and international concordats and guidance, such as the *Transparency and Openness Promotion (TOP) Guidelines* (Centre for Open Science, 2025), the *UK Concordat on Open Research Data* (The Concordat Working Group, 2016), *NIH 2023 Data Management & Sharing policy* (NIH, 2023), *the European Code of Conduct on Research Integrity* (ALLEA, 2023). Indirectly, institutional policies may contribute to mitigating self-reported engagement in QRPs; however, the extent of their actual impact on researcher behaviour remains relatively understudied and lacks substantial empirical investigation.

Individual-level Factors

Attributing researcher misbehaviour solely to asymmetric incentive structures is both reductionist and deterministic. Although a growing body of evidence – and significant theory – links misaligned incentives to sub-optimal research practices, treating such structures as the sole causal mechanism is analytically convenient and obscures the social, cognitive, and psychological determinants that also shape scientific conduct. Correlational research suggests that individual-level factors may be as important as *systemic-level* pressures in predicting self-reported QRP engagement (Schneider et al., 2024; Gopalakrishna et al., 2022). For example, cognitive biases when navigating research decisions (John et al., 2012; MacCoun, 1998; Nuzzo, 2015) and competency shortfalls in methodology (John et al., 2012; Munafò et al., 2017), data analysis (MacCoun, 2022) and research ethics (Johnsson et al., 2014; Resnik, 2011) have been found to influence researcher practice.

The proposition that deficits in methodological competence underlie low-quality, irreproducible, and ethically questionable research findings strengthens the case for a heightened emphasis on rigorous integrity and ethics training. This awareness has translated into structured training in integrity being offered by many institutions, sometimes as core curricular for doctoral training, and is favoured as a fix for suboptimal research practice. Indeed, some academic organisations, such as the University of Oxford (University of Oxford, 2023), and funders such as the NIH and the National Science Foundation, mandate integrity training as a condition of study/funding. However, a study of biomedical researchers in the Netherlands finds that childhood education and personality traits may have a greater influence on research practice than formal training (Satalkar & Shaw, 2019). Moreover, the limited number of studies examining the relationship between research integrity training and actual researcher behaviour have generally reported either negligible or weak associations (Anderson et al., 2007b; Antes et al., 2010). On this basis, there is debate as to what extent integrity training is a remedy for engagement in QRPs. That being said, there is likely nuance to the relationship. For example, studies have shown that researchers who strive for skill development engage in QRPs significantly less than those who strive for skill demonstration (Janke et al., 2019). On this basis, such tests of efficacy of integrity training are likely confounded by a number of factors, including researcher motivation. Studies investigating the effectiveness of integrity training should take such nuance into account.

Beyond this, limited research has sought to empirically investigate the roles of researchers' personal motivations or science-specific motivations on researcher behaviour. Studies into individual differences have found that agreeableness and conscientiousness, are negatively correlated with self-reported QRP engagement (Schneider et al., 2024; Giluk &

Postlethwaite, 2015). Moreover, it has been found that Machiavellianism is positively correlated with self-reported QRP engagement (Tijdink et al., 2016). However, empirical research examining the psychological determinants of engagement in QRPs remains limited.

A commitment to normative ideals that are conducive with good science may prevent QRP engagement, especially in the face of systemic incentives. Robert Merton in *The Normative Structure of science* (1942) delineated four normative ideals, which act as "prescriptions, proscriptions, preferences, and permissions" for researchers within the ecosystem of science. These values, according to Merton, are internalised by researchers and become a part of their conscience. These normative ideals, include communalism (scientific discoveries are common heritage), universalism (truth-claims should be evaluated by impersonal, pre-established criteria), disinterestedness (personal or commercial gain should be subordinate to objective inquiry), and organised scepticism (the submission of claims to systematic, organised criticism). These normative ideals function to keep researchers motivated to the collective goal of extending certified knowledge. There has been ongoing debate about the extent to which these values are internalised by researchers (Mitroff, 1974; Macfarlane, 2023; Koppel et al., 2025) and the degree to which they shape everyday scientific conduct. Even Merton in his original delineation stated that internalisation of these norms by researchers occurs to varying degrees, and that social, political, and economic forces may pull researchers in the opposite direction. Moreover, Mulkay (1976) argued that these norms are an 'ideology' that researchers use to legitimise their craft and are not a reliable guide to actual behaviour.

Empirical evidence, primarily from self-report surveys, shows that internalisation of these normative ideals is widespread, but not universal (Anderson et al., 2010; Bray & von Storch, 2017; Macfarlane & Cheng, 2008; Kim & Kim, 2018). To date, the small body of empirical work examining the link between scientific norms and poor scientific conduct has relied almost exclusively on correlational studies. One study, based on a survey of researchers in the Netherlands, found that a stronger commitment to the normative values of science (operationalised as a single construct) was positively correlated with self-reported QRP engagement (Gopalakrishna et al., 2022).

Moreover, the extent to which the benefits of engaging in QRPs are perceived to outweigh the potential consequences ('risks'), the greater the self-reported engagement (Gopalakrishna et al., 2022; Bruton et al., 2019). A study of researchers in the Netherlands found that perceived likelihood of QRP detection by reviewers correlated with reduced self-reported engagement in these practices (Gopalakrishna et al., 2022). Relatedly, Bruton et al. (2019) found that among a sample of NIH-funded scientists, where QRPs are deemed as riskier, they were less willing to engage in them. In this sense, academia can be construed as producing a criminogenic opportunity structure, whereby academics are motivated offenders, with systemic incentives (publication, tenure, promotion) motivating researchers to engage in such behaviours, with low chances of detection and sanction increasing the likelihood of engagement. This presents another potential avenue of remedy, where bolstering detection mechanisms (e.g., with open science practices, such as pre-registration, mandating data and code availability) and maximising real-world consequences (greater sanction) for transgressing integrity standards may offer crucial levers for fostering research integrity.

Finally, the social-psychological process of normative influence feasibly has an effect on how researchers behave. Descriptive norms – people's perceptions of what "most others" do (Bicchieri, 2017) – can encourage unethical academic behaviour, including academic cheating (McCabe, 2002; Smyth & Davis, 2004; Chapman et al., 2004; Stone et al., 2009) and scientific misconduct (Sacco et al., 2018; Rajah-Kanagasabai & Roberts, 2015). Conversely, injunctive norms – desirable behaviours (Cialdini et al., 1990) – may influence behaviour by indicating what behaviours are ethically acceptable. It is feasible that these normative processes influence behaviour in the direction of the norm in the case of more benign and less severe cases of research misbehaviour, including QRPs. Perceiving QRPs to be descriptively normative may motivate researchers to engage in these behaviours through processes of normative influence (Asch, 1956; Sherif, 1936) and informational social influence (Deutsch & Gerrard, 1955).

Paper Synthesis

In this doctoral thesis I respond to calls made elsewhere for the identification of the factors and mechanisms that facilitate and motivate engagement in dishonest scientific practices (Paruzel-Czachura et al., 2020). Existing literature point towards a multi-level aetiology: misaligned structures that incentivise corner-cutting, institutional cultures and governance that amplify or dampen that risk, and individual-level factors (orientations, values, motivations) that determine whether researchers cross the line. The evidence base is heavily fragmented and fails to model cross-level interactions simultaneously. The three, separate empirical papers contained in this thesis contribute to an integrated conceptual model that delineates how systemic incentives, institutional climates, perceived social norms, and individual dispositions interact to drive engagement in QRPs. Across three empirical papers, the following questions are addressed:

1. How much of the variance in self-reported QRP engagement lies at the institution and country levels, and which contextual factors explain it? (Paper A)
2. Are researchers who think their peers engage in QRPs more likely to report doing so themselves? (Paper B)
3. In an integrated model, what are the relative associations of individual-level and systemic incentives in predicting QRP use? (Paper C)

The first empirical paper (Paper A) develops an explanatory model by investigating the correlates of individual-, organisational- and systemic level factors on self-reported engagement in QRPs. Further to this, it seeks to recognise the relative contribution of individual-level and organisational-level factors on self-reported QRP engagement. If there are significant organisational-level differences in QRP self-reported engagement, there is an impetus for empirical investigation to focus on institutional-level factors that correlate with self-reported engagement, and additionally institutional-targeted solutions (i.e., adjustments to governance, culture, and policy). On the contrary, if institutional variance in self-reported QRP engagement is minimal, individual-level, and system-wide, factors should the primary targets of investigation and intervention. To date, no study has examined variability in self-reported QRP engagement across universities, with pooled analyses being standard practice. On this basis, we use multilevel modelling to analyse data from the International Survey on Research Integrity (IRIS; Allum et al., 2023). This secondary data contains responses to a batch of questions from researchers operating across disciplinary fields and countries. Specifically, it investigates whether individual-level factors (such as commitment to Mertonian norms of science, disciplinary field, contract type, career stage and sex) and organisation level factors (such as, workplace type, presence of integrity training and institutional statements on research integrity and organisational culture) correlate with the extent of self-reported QRP engagement.

Preliminary evidence suggests that researchers' perceptions of their colleagues' practices correlate with their own self-reported engagement in QRPs, aligning behaviour with perceived norms. For instance, Sacco et al. (2017) found that principal investigators funded by the NIH in the United States were more likely to consider QRPs as ethically acceptable when such practices were perceived as normative among their peers. This effect likely operates through mechanisms of competitive pressure (i.e., a desire to 'keep up' with peers) or through normative judgements about ethical acceptability (i.e., 'my peers engage in this practice, therefore it must be permissible'). On this basis, the second empirical paper (Paper B) investigates a) whether perceptions of normativity correlate with self-reported QRP engagement, and b) the mechanisms through which this occurs. If perceptions are a key driver of self-reported QRP engagement, interventions can target misperceived norms, rather than individual morality or competency. It opens the door for possible leverage points using novel interventional strategies grounded in behavioural economics and social psychology, to reduce self-reported QRP engagement. Paper B utilises data from the Practices, Perceptions, and Patterns of Research Integrity' (PRINT) survey of researchers operating across disciplinary fields in the US and Europe. The paper uses Bayesian multi-level modelling to regress self-reported QRP engagement on perceived prevalence, while analysing the mechanisms through which this process occurs.

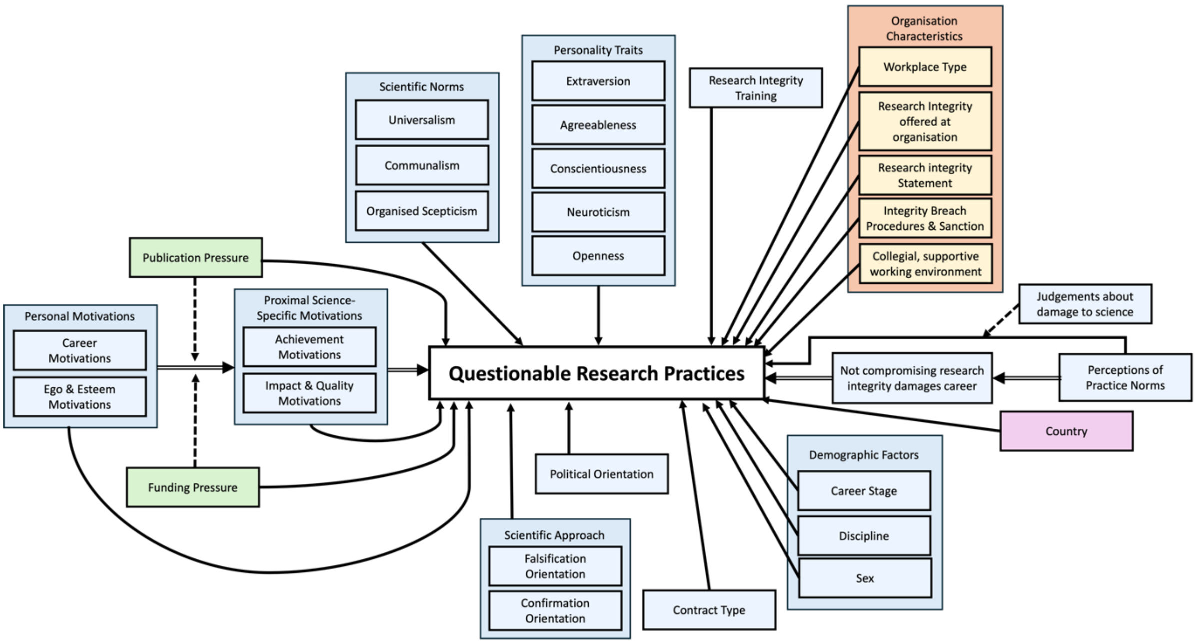
The final paper (Paper C) tests an integrated model of the possible correlates of self-reported QRP engagement. It builds on Papers A and B by focusing on potential individual-level and system-level factors of self-reported QRP engagement. This final paper seeks to present a coherent and integrative empirical model, and partially builds on the theoretical model presented by Jussim et al. (2019), which outlines how researcher competencies, motivations, system-level factors and social-psychological influences – such as perceptions of practice norms – correlate with self-reported engagement.[4] Specifically, Paper C investigates how personal motivations (career advancement, ego, and esteem), proximal science-specific motivations (achievement, impact, and quality), commitment to scientific norms, personality traits, political orientation, and epistemic orientation (falsification versus confirmation) correlate with researchers' self-reported engagement in QRPs. Additionally, the paper analyses the relative importance of systemic pressures – including funding and publication pressures – and explores the interplay between these individual-level and system-level factors in predicting self-reported engagement in QRPs. This comprehensive analysis identifies the leverage points

[4] Jussim et al.'s (2019) Social Psychological Model of Scientific Practices provides a comprehensive (though, not exhaustive) theoretical framework of the possible determinants, but also downstream consequences, of potentially sub-optimal research practices.

with the highest payoff of reform. The paper uses primary survey data, collected from a sample of researchers operating across disciplinary fields and countries.

Collectively, these papers move the explanation for self-reported QRP engagement from organisational context, through individual determinants, to a full multilevel account. Figure 1.1 presents our proposed theoretical model and serves as a visual illustration of our hypotheses and exploratory analysis across the three papers.[5]

Figure 1.1 Conceptual model illustrating the individual-, organisational- and system-level determinants of QRPs



Each paper is written in the format of a standalone paper, as they are intended for publication in an academic journal (and, in one case, have already been published).

Data and Methodological Approach

Paper A (Brooker & Allum 2024; Chapter 2) uses secondary data collected as part of the Standard Operating Procedures for Research Integrity (SOPs4RI) project in the

[5]Figure 1 should not be construed as an overview of the analytical strategy, but instead a visualisation of the conceptual framework for investigation.

International Survey on Research Integrity (IRIS). The survey was administered to a sample of 60,000 researchers operating across disciplinary fields primarily in Europe and the United States. The study seeks to investigate the relative impact of individual-level and organisational-level factors in determining self-reported QRP engagement. It also seeks to identify the extent self-reported to which QRP engagement varies across countries, thus addressing whether national context is important. To achieve this, a multilevel model will be employed to estimate the proportion of variance attributable to each level of the data hierarchy. This is exploratory in nature. In addition to identifying the country-level, institutional-level and individual-level variance, across six multilevel linear regression models, fixed effects at both the individual- and institutional-levels are estimated. The individual-level fixed effects include contract type, career stage, disciplinary field, sex, and commitment to scientific norms, while the organisational-level fixed effects include workplace type (e.g., industry, non-profit, academia), awareness of an institutional research integrity statement, collegiality and competitiveness of the working environment, procedures for addressing integrity breaches and presence of integrity training. The degree to which researchers are committed to the normative ideals of science, operationalised as a single, unified construct, and the relationship of this commitment with self-reported QRP engagement, is central to this paper. I additionally estimate cross-level interactions.

Paper B (Chapter 3) was pre-registered on the open science framework (OSF) on 5th December 2021 (Available at: https://shorturl.at/QuqM2). The paper analyses data from the Practices, Perceptions and Patterns of Research Integrity (PRINT) survey, administered to a sample of 55,000 researchers operating across disciplinary fields in Denmark, Croatia, the United Kingdom, Austria, and the United States. The survey, fielded in 2018, gathered data on QRP prevalence, perceptions of prevalence, reward structures and pressure, and personality traits. The final analytical sample consisted of 4,646 researchers representing all major disciplines. The study tests a series of hypotheses:

H1: The more prevalent a researcher perceives QRPs to be within their scientific field, the greater the extent of their own use of those practices.

H2: Estimates of QRP prevalence within a researcher's field are indirectly associated with self-reported engagement in those practices, through beliefs that maintaining research integrity may have negative career consequences.

H3: The relationship between perceptions of prevalence of QRPs within a researcher's field and the extent of their own use of those practices will be greater when the practices are seen as more damaging to science.

H4: The relationship between perceptions of prevalence of QRPs within a researcher's scientific field and the degree to which the researcher has used those practices varies according to the researchers' career stage.

H5: Researchers perceive the prevalence of QRPs to be greater in their scientific field compared to in their own sole or co-authored publications.

The study employs a Bayesian multilevel modelling framework to account for both the non-independence of responses to QRP items within individual respondents and the hierarchical structure of the data. The Bayesian framework permits the derivation of a posterior distribution that expresses the uncertainty in the parameter values, providing a natural quantification of the uncertainty in inferences. Using multilevel modelling allows us to control for dependencies between numerous responses to QRP items within the same respondent, while allowing us to measure how perceived prevalence of each QRP is associated with self-reported engagement in *that* particular QRP. That is, it permits us to investigate measurement-level variability – variability that is lost when average scores are computed across QRP items. I regress self-reported QRP engagement on estimates of QRP prevalence using muti-level ordered categorical models. It additionally models how measurement level variables, such as judgements about the damage these practices have on science, and respondent-level variables, including discipline and career stage, relate to self-reported QRP engagement.

Paper C (Chapter 4) analyses the data from a primary cross-sectional survey of researchers operating across disciplinary fields and countries. The sample is drawn from individuals who, via the IRIS platform, indicated their willingness to be contacted for participation in future research. The survey was administered between November and December 2024 and comprises items that assess researcher behaviour, scientific misconduct, research culture, normative ideals, motivations, attitudes toward academic freedom, self-censorship, and demographic characteristics. The study was pre-registered on OSF in November 2024 (available at: https://shorturl.at/mf9qY) and was administered to a sample of 18,376 researchers. The final analytical sample consisted of 3,109 researchers. In Paper C, I test the following hypotheses:

H1: The more individuals are motivated by career achievement, the greater their self-reported QRP engagement.

H2: The more individuals are motivated by ego and esteem, the greater their self-reported QRP engagement.

H3: Scientific achievement (e.g., obtaining a higher citation index, publishing more frequently) is positively associated with self-reported QRP engagement.

H4: A motivation to increase the impact and quality of one's research is negatively associated with self-reported QRP engagement.

H5-H8: Proximal science-specific motivations mediate the relationship between personal motivations and self-reported QRP engagement.

H9: Self-reported publication pressure is positively associated with self-reported QRP engagement.

H10: Funding pressure is positively associated with self-reported QRP engagement.

H11-H14: Publication pressure moderates the relationship between personal motivations and proximal science-specific motivations.

H15-H18: Funding pressure moderates the relationship between personal motivations and proximal science-specific motivations.

H19a: Conscientiousness is negatively associated with self-reported QRP engagement.

H19b: Agreeableness is negatively associated with self-reported QRP engagement.

The study includes additional exploratory elements. Firstly, it explores whether a researcher's epistemic orientation (the extent to which a researcher seeks to confirm versus falsify one's theories or findings) correlates with the extent to which they engage in QRPs. Secondly, it seeks to investigate whether political orientation correlates with self-reported QRP engagement. Thirdly, it explores whether a commitment to the Mertonian norms of science correlates with self-reported QRP engagement. However, unlike in Paper A (Chapter 2), it operationalises Mertonian norms as distinct constructs and investigates their *relative* and *differential* impact on QRPs. Across a series of linear regression and logistic regression models, the primary explanatory variables of interest (publication and funding pressure, personal motivations, proximal science-specific motivations, epistemic orientation, personality traits and scientific norms) are regressed on mean QRP and a binary variable indicating whether an individual has engaged in at least one QRP frequently. Interaction terms are included, seeking to investigate whether funding and publication pressure moderate the relationship between personal motivations and proximal science-specific motivations, as well as between scientific norms and self-reported QRP engagement. Moreover, I include interaction terms to explore whether the relationship between personality traits and self-reported QRP engagement are moderated by publication pressure. Finally, Structural Equation Modelling (SEM) is used to test whether proximal science-specific motivations mediate the relationship between personal motivations and QRPs.

Table 1.1. Summary of Research Design, Data Sources, and Analytical Methods Across Studies

Primary Research Question	Dependent Variable	Explanatory Factors	Statistical Analysis	Dataset	Analytical Sample	Paper / Chapter
How much of the variance in self-reported QRP engagement lies at the institution and country levels, and which contextual factors explain it?	**Mean QRP** (based on eight QRP items)	**Individual level:** normative ideals of science, contract type, career stage, disciplinary field, sex. **Organisational level:** working environment, workplace type, integrity breaches, integrity training and RI statement awareness. **Random Effects:** Organisation and Country.	Frequentist Multi-level Modelling (Multiple Linear Regression)	International Survey on Research Integrity (secondary dataset)	39,699 researchers operating across disciplinary fields and countries.	Paper A / Chapter 2
Are researchers who think their peers engage in QRPs more likely to report doing so themselves?	**Measurement-level** self-reported **QRP engagement.**	**Measurement level:** Estimates of QRP prevalence; damage to trust in scientific findings. **Individual level:** Gender; country; research integrity training; career stage; disciplinary field; career necessity.	Bayesian Multi-level Modelling (Logit Regression)	International survey fielded as part of the Practices, Perceptions and Patterns of Research Integrity (PRINT) project	4646 researchers across fields in Denmark, Croatia, the UK, Austria and the United States.	Paper B / Chapter 3
In an integrated model, what are the relative correlates of individual-level and systemic incentives in predicting QRP use?	▪ **Mean QRP** (looped, determined by methodological approach). ▪ **Any frequent QRP** (dichotomous, binary variable)	**Individual -level:** Personal motivations (career motivations, ego & esteem); Proximal science-specific motivations (achievement, impact & quality); Scientific norms (universalism, communalism, organised scepticism); Personality traits (extraversion, agreeableness, conscientiousness, neuroticism, openness); Political orientation; Scientific approach (falsification, confirmation. **System-level:** Publication pressure; Funding pressure.	▪ Factor analysis. ▪ Linear & logistic regression models ▪ Structural equation modelling (SEM)	Primary cross-sectional survey of researchers operating across disciplinary fields and countries.	3109 researchers working across fields and countries.	Paper C / Chapter 4

Anticipated Contributions

Collectively, the three studies included in this thesis convert scattered speculation, and fragmented empirical inquiry, into a comparative, quantified evidence base. Paper A shows the extent to which variability in self-reported QRP engagement can be attributed to country (systemic), organisational and individual level differences. Paper B demonstrates the extent to which perceiving behaviours to be descriptively normative correlates with increased self-reported QRP engagement, as well as elucidating the mechanisms through which this may occur. Paper C investigates whether personal and proximal-science specific motivations, a commitment to scientific norms, personality traits and epistemic orientation are *as important* predictors of self-reported QRP engagement as systemic factors, such as publication and funding pressure.

Together, these studies offer a theoretically grounded and empirically rich account of the multilevel drivers of QRPs. This thesis advances understanding in three key ways. Firstly, it provides one of the few comparative, large-scale examinations of self-reported QRP engagement across national and institutional contexts, clarifying the relative weight of organisational versus individual-level correlates. Secondly, it contributes novel insights into the role of social norms, highlighting the behavioural mechanisms through which perceptions of peer practice shape researchers' own practices. Thirdly, it integrates psychological, normative, and systemic predictors within a unified analytical framework, demonstrating that individual motivations and epistemic commitments are not merely ancillary but central to understanding questionable research practice. By weaving together these dimensions, the thesis moves the literature beyond isolated explanations and toward a more comprehensive account of why researchers compromise scientific integrity.

Given the detrimental downstream consequences of QRPs – both for the integrity of the scientific record and the effectiveness of evidence-based interventions – this thesis contributes to identifying key leverage points for reducing engagement. By illuminating the conditions under which QRPs are more likely to occur, and the factors that may buffer against them, it provides an empirically grounded basis for developing targeted reforms in research governance, training, and institutional culture.

References

Adda, J., Decker, C. and Ottaviani, M. (2020) ‘p-hacking in clinical trials and how incentives shape the distribution of results across phases’, *Proceedings of the National Academy of Sciences*, 117(24), pp. 13386–13392. Available at: https://doi.org/10.1073/pnas.1919906117 (Accessed: 3 June 2026).

Ai, F., Zhao, J., Yang, W., & Wan, X. (2023). Dose-response relationship between active smoking and lung cancer mortality/prevalence in the Chinese population: a meta-analysis. *BMC public health*, *23*(1), 747. https://doi.org/10.1186/s12889-023-15529-7

ALLEA. (2023). *European Code of Conduct for Research Integrity*. Revised Edition. https://allea.org/wp-content/uploads/2023/06/European-Code-of-Conduct-Revised-Edition-2023.pdf

Allum, N., Reid, A., Bidoglia, M., Gaskell, G., Aubert-Bonn, N., Buljan, I., Fuglsang, S., Horbach, S., Kavouras, P., Marušić, A., Mejlgaard, N., Pizzolato, D., Roje, R., Tijdink, J., & Veltri, G. (2023). Researchers on research integrity: a survey of European and American researchers. *F1000Research*, *12*, 187. https://doi.org/10.12688/f1000research.128733.1

Anderson, M. S., Horn, A. S., Risbey, K. R., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007b). What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a National Survey of NIH-funded scientists. Academic medicine : journal of the Association of American Medical Colleges, 82(9), 853–860. https://doi.org/10.1097/ACM.0b013e31812f764c

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007a). The perverse effects of competition on scientists' work and relationships. *Science and engineering ethics*, *13*(4), 437–461. https://doi.org/10.1007/s11948-007-9042-5

Anderson, M.S., Ronning, E.A., De Vries, R., & Martinson, B.C. (2010). Extending the Mertonian Norms: Scientists' Subscription to Norms of Research. *The Journal of Higher Education 81*(3), 366-393. https://dx.doi.org/10.1353/jhe.0.0095.

Antes, A. L., Wang, X., Mumford, M. D., Brown, R. P., Connelly, S., & Devenport, L. D. (2010). Evaluating the effects that existing instruction on responsible conduct of research has on ethical decision making. *Academic medicine : journal of the Association of American Medical Colleges*, *85*(3), 519–526. https://doi.org/10.1097/ACM.0b013e3181cd1cc5

Armond, A. C. V., & Kakuk, P. (2022). Perceptions of Research Integrity Climate in Hungarian Universities: Results from A Survey among Academic Researchers. *Science and engineering ethics*, *28*(4), 30. https://doi.org/10.1007/s11948-022-00382-5

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied, 70*(9), 1–70. https://doi.org/10.1037/h0093718

Baker M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452a

Bartneck, C., Kokkelmans, S. (2010). Detecting *h*-index manipulation through self-citation analysis. *Scientometrics 87*(1), 85–98 https://doi.org/10.1007/s11192-010-0306-5

Bicchieri, C. (2017). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York, online ed, Oxford Academic. https://doi.org/10.1093/acprof:oso/9780190622046.001.0001,

Bik, E. M., Casadevall, A., & Fang, F. C. (2016). The Prevalence of Inappropriate Image Duplication in Biomedical Research Publications. *mBio*, *7*(3), e00809-16. https://doi.org/10.1128/mBio.00809-16

Bordignon, F., Chaignon, L., & Egret, D. (2023). Promoting narrative CVs to improve research evaluation? A review of opinion pieces and experiments. *Research Evaluation*, *32*(2), 313–320. https://doi.org/10.1093/reseval/rvad013

Bouter L. M. (2015). Commentary: Perverse incentives or rotten apples?. *Accountability in research*, *22*(3), 148–161. https://doi.org/10.1080/08989621.2014.950253

Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & Ter Riet, G. (2019). Correction to: ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. Research integrity and peer review, 4, 13. https://doi.org/10.1186/s41073-019-0072-8

Bray, D., & von Storch, H. (2017). The Normative Orientations of Climate Scientists. *Science and engineering ethics*, *23*(5), 1351–1367. https://doi.org/10.1007/s11948-014-9605-1

Bruton, S. V., Brown, M., Sacco, D. F., & Didlake, R. (2019). Testing an active intervention to deter researchers' use of questionable research practices. *Research integrity and peer review*, *4*, 24. https://doi.org/10.1186/s41073-019-0085-3

Bucci E. M. (2018). Automatic detection of image manipulations in the biomedical literature. *Cell death & disease*, *9*(3), 400. https://doi.org/10.1038/s41419-018-0430-3

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., … Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. Nature Human Behaviour, 2(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Center for Open Science. (2025). *The TOP Guidelines were created by journals, funders, and societies to align scientific ideals with practices*. https://www.cos.io/initiatives/top-guidelines

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Lancet (London, England)*, *374*(9683), 86–89. https://doi.org/10.1016/S0140-6736(09)60329-9

Chapman, K. J., Davis, R., Toy, D., & Wright, L. (2004). Academic Integrity in the Business School Environment: I'll Get by with a Little Help from My Friends. *Journal of Marketing Education*, *26*(3), 236-249. https://doi.org/10.1177/0273475304268779

Chavalarias, D., Wallach, J. D., Li, A. H., & Ioannidis, J. P. (2016). Evolution of Reporting P Values in the Biomedical Literature, 1990-2015. *JAMA*, *315*(11), 1141–1148. https://doi.org/10.1001/jama.2016.1952

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Crain, A. L., Martinson, B. C., & Thrush, C. R. (2013). Relationships between the Survey of Organizational Research Climate (SORC) and self-reported research practices. *Science and engineering ethics*, *19*(3), 835–850. https://doi.org/10.1007/s11948-012-9409-0

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgement. *Journal of abnormal psychology*, *51*(3), 629–636. https://doi.org/10.1037/h0046408

Dwan, K., Gamble, C., Williamson, P. R., Kirkham, J. J., & Reporting Bias Group (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PloS one*, *8*(7), e66844. https://doi.org/10.1371/journal.pone.0066844

Edwards, M. A., & Roy, S. (2017). Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental engineering science*, *34*(1), 51–61. https://doi.org/10.1089/ees.2016.0223

Erasmus, A. (2024) 'p-hacking: Its costs and when it is warranted', *Erkenntnis*. Available at: https://doi.org/10.1007/s10670-024-00834-3 (Accessed: 3 June 2026).

Fanelli D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PloS one, 4(5), e5738. https://doi.org/10.1371/journal.pone.0005738

Fanelli, D. (2018) 'Is science really facing a reproducibility crisis, and do we need it to?', *Proceedings of the National Academy of Sciences*, 115(11), pp. 2628–2631. Available at: https://doi.org/10.1073/pnas.1708272114 (Accessed: 3 June 2026).

Fiedler, K., & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological and Personality Science, 7*(1), 45–52. https://doi.org/10.1177/1948550615612150

Fire, M., & Guestrin, C. (2019). Over-optimization of academic publishing metrics: observing Goodhart's Law in action. *GigaScience*, *8*(6), giz053. https://doi.org/10.1093/gigascience/giz053

Francis, G. (2012) 'Publication bias and the failure of replication in experimental psychology', *Psychonomic Bulletin & Review*, 19(6), pp. 975–991. Available at: https://doi.org/10.3758/s13423-012-0322-y (Accessed: 3 June 2026).

Gandini, S., Botteri, E., Iodice, S., Boniol, M., Lowenfels, A. B., Maisonneuve, P., & Boyle, P. (2008). Tobacco smoking and cancer: a meta-analysis. *International journal of cancer*, *122*(1), 155–164. https://doi.org/10.1002/ijc.23033

Giluk, T. L., & Postlethwaite, B. E. (2015). Big five personality and academic dishonesty: A meta-analytic review. *Personality and Individual Differences, 72,* 59–67. https://doi.org/10.1016/j.paid.2014.08.027

Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PloS one*, *17*(2), e0263023. https://doi.org/10.1371/journal.pone.0263023

Gupta, A. and Bosco, F. (2023) 'Tempest in a teacup: an analysis of p-hacking in organizational research', *PLOS ONE*, 18(2), e0281938. Available at: https://doi.org/10.1371/journal.pone.0281938 (Accessed: 3 June 2026).

Haven, T. L., Tijdink, J. K., Martinson, B. C., & Bouter, L. M. (2019). Perceptions of research integrity climate differ between academic ranks and disciplinary fields: Results from a survey among academic researchers in Amsterdam. *PloS one*, *14*(1). https://doi.org/10.1371/journal.pone.0210599

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS biology*, *13*(3), e1002106. https://doi.org/10.1371/journal.pbio.1002106

Hendriks, F., Kienhues, D., & Bromme, R. (2020). Replication crisis = trust crisis? The effect of successful vs failed replications on laypeople's trust in researchers and research. *Public understanding of science (Bristol, England)*, *29*(3), 270–288. https://doi.org/10.1177/0963662520902383

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429–431. https://doi.org/10.1038/520429a

Horie, N., & Iwasaki, I. (2022). Returns to schooling in European emerging markets: a meta-analysis. *Education Economics*, *31*(1), 102–128. https://doi.org/10.1080/09645292.2022.2036322

Hwang, S. Y., Yon, D. K., Lee, S. W., Kim, M. S., Kim, J. Y., Smith, L., Koyanagi, A., Solmi, M., Carvalho, A. F., Kim, E., Shin, J. I., & Ioannidis, J. P. A. (2023). Causes for Retraction in the Biomedical Literature: A Systematic Review of Studies of Retraction Notices. *Journal of Korean medical science*, *38*(41), e333. https://doi.org/10.3346/jkms.2023.38.e333

Ioannidis J. P. (2005). Why most published research findings are false. *PLoS medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Janke, S., Daumiller, M., & Rudert, S. C. (2019). Dark pathways to achievement in science: Researchers' achievement goals predict engagement in questionable research practices. *Social Psychological and Personality Science, 10*(6), 783–791. https://doi.org/10.1177/1948550618790227

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. https://doi.org/10.1177/0956797611430953

Johnson, L., Erikson, S., Helgesson, G., & Hansson, M. G. (2014). Making researchers moral: Why trustworthiness requires more than ethics guidelines and review. *Research Ethics*, *10*(1), 29-46. https://doi.org/10.1177/1747016113504778

Jussim, L., Krosnick, J. A., Stevens, S. T., & Anglin, S. M. (2019). A Social Psychological Model of Scientific Practices: Explaining Research Practices and Outlining the Potential for Successful Reforms. *Psychologica Belgica*, *59*(1), 353–372. https://doi.org/10.5334/pb.496

Kerr N. L. (1998). HARKing: hypothesizing after the results are known. *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc*, *2*(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4

Kim, S. Y., & Kim, Y. (2018). The Ethos of Science and Its Correlates: An Empirical Analysis of Scientists' Endorsement of Mertonian Norms. *Science, Technology and Society*, *23*(1), 1-24. https://doi.org/10.1177/0971721817744438

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Jr., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., ... Nosek, B. A. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443–490. https://doi.org/10.1177/2515245918810225

Koppel, L., Lindkvist, A. M., & Tinghög, G. (2025). Scientific normative dissonance: A large-scale survey of researchers' subscription to scientific norms and counternorms across academic fields. *Journal of Economic Behavior & Organization*, *237*(107140). https://doi.org/10.1016/j.jebo.2025.107140

Kretser, A., Murphy, D., Bertuzzi, S., Abraham, T., Allison, D. B., Boor, K. J., Dwyer, J., Grantham, A., Harris, L. J., Hollander, R., Jacobs-Young, C., Rovito, S., Vafiadis, D., Woteki, C., Wyndham, J., & Yada, R. (2019). Scientific Integrity Principles and Best Practices: Recommendations from a Scientific Integrity Consortium. *Science and engineering ethics*, *25*(2), 327–355. https://doi.org/10.1007/s11948-019-00094-3

Lee, P.N., Forey, B.A. & Coombs, K.J. (2012). Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC Cancer* 12, 385 https://doi.org/10.1186/1471-2407-12-385

Leung, K. (2011) 'Presenting post hoc hypotheses as a priori: ethical and theoretical issues', *Management and Organization Review*, 7(3), pp. 471–479. Available at: https://doi.org/10.1111/j.1740-8784.2011.00222.x (Accessed: 3 June 2026).

MacCoun R. J. (1998). Biases in the interpretation and use of research results. *Annual review of psychology*, *49*, 259–287. https://doi.org/10.1146/annurev.psych.49.1.259

MacCoun, R. J. (2022). P-hacking: A strategic analysis. In L. J. Jussim, J. A. Krosnick, & S. T. Stevens (Eds.), *Research integrity: Best practices for the social and behavioral sciences* (pp. 295–313). Oxford University Press. https://doi.org/10.1093/oso/9780190938550.003.0011

Macfarlane, B. (2023). The DECAY of Merton's scientific norms and the new academic ethos. *Oxford Review of Education*, *50*(4), 468–483. https://doi.org/10.1080/03054985.2023.2243814

Macfarlane, B., Cheng, M. (2008). Communism, Universalism and Disinterestedness: Re-examining Contemporary Support among Academics for Merton's Scientific Norms. *Journal of Academic Ethics* *6*(1), 67–78. https://doi.org/10.1007/s10805-008-9055-y

MacLeod C. M. (1991). Half a century of research on the Stroop effect: an integrative review. Psychological bulletin, 109(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Maggio, L., Dong, T., Driessen, E., & Artino, A., Jr (2019). Factors associated with scientific misconduct and questionable research practices in health professions education. *Perspectives on medical education*, *8*(2), 74–82. https://doi.org/10.1007/s40037-019-0501-x

Martinson, B. C., Anderson, M. S., & de Vries, R. (2005). Scientists behaving badly. *Nature,* *435*(7043), 737–738. https://doi.org/10.1038/435737a

Martinson, B. C., Anderson, M. S., Crain, A. L., & de Vries, R. (2006). Scientists' perceptions of organizational justice and self-reported misbehaviors. *Journal of empirical research on human research ethics : JERHRE*, *1*(1), 51–66. https://doi.org/10.1525/jer.2006.1.1.51

McCabe, D. L., Treviño, L. K., & Butterfield, K. D. (2001). Cheating in academic institutions: A decade of research. *Ethics & Behavior, 11*(3), 219. https://doi.org/10.1207/S15327019EB1103_2

Moher D. (2018). Reporting guidelines: doing better for readers. *BMC medicine*, *16*(1), 233. https://doi.org/10.1186/s12916-018-1226-0

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., Coriat, A. M., Foeger, N., & Dirnagl, U. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLoS biology*, *18*(7), e3000737. https://doi.org/10.1371/journal.pbio.3000737

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS biology*, *16*(3), e2004089. https://doi.org/10.1371/journal.pbio.2004089

Mohseni, A. (2020) *HARKing: from misdiagnosis to misprescription*. PhilSci-Archive. Available at: https://philsci-archive.pitt.edu/18523/ (Accessed: 3 June 2026).

Montenegro, C. E., & Patrinos, H. A. (2014). Comparable Estimates of Returns to Schooling Around the World. *World Bank: Policy Research Working Paper 7020*. https://doi.org/10.1596/1813-9450-7020.

Mulkay, M. (1976). Norms and ideology in science. *Social Science Information, 15*, 637 - 656. https://doi.org/10.1177/053901847601500406

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature human behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

NIH. (2023). *Data Management & Sharing Policy Overview*. https://grants.nih.gov/policy-and-compliance/policy-topics/sharing-policies/dms/policy-overview#after

Norman G. (2014). Data dredging, salami-slicing, and other successful strategies to ensure rejection: twelve tips on how to not get your paper published. *Advances in health sciences education : theory and practice*, *19*(1), 1–5. https://doi.org/10.1007/s10459-014-9494-8

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on psychological science : a journal of the Association for Psychological Science*, *7*(6), 615–631. https://doi.org/10.1177/1745691612459058

Nuzzo R. (2015). How scientists fool themselves - and how they can stop. *Nature*, *526*(7572), 182–185. https://doi.org/10.1038/526182a

O'Keeffe, L. M., Taylor, G., Huxley, R. R., Mitchell, P., Woodward, M., & Peters, S. A. E. (2018). Smoking as a risk factor for lung cancer in women and men: a systematic review and meta-analysis. BMJ open, 8(10), e021611. https://doi.org/10.1136/bmjopen-2018-021611

Oberauer, K. and Lewandowsky, S. (2019) ‘Addressing the theory crisis in psychology’, *Psychonomic Bulletin & Review*, 26(5), pp. 1596–1618. Available at: https://doi.org/10.3758/s13423-019-01645-2 (Accessed: 3 June 2026).

Office for National Statistics. (2024, October 1). *Adult smoking habits in the UK: 2023*. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/adultsmokinghabitsingreatbritain/2023#:~:text=The%20

Annual%20Population%20Survey%20(APS)%20gives%20headline%20indicators%20on%20the,Scotland%20was%2013.5%25

Open Science Collaboration (2015). PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science (New York, N.Y.)*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Paruzel-Czachura, M., Baran, L., & Spendel, Z. (2020). Publish or be ethical? Publishing pressure and scientific misconduct in research. *Research Ethics*, *17*(3), 375-397. https://doi.org/10.1177/1747016120980562

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science, 7*(6), 531–536. https://doi.org/10.1177/1745691612463401

Post A, Li A Y, Dai J B, et al. (2018) c-index and Subindices of the h-index: New Variants of the h-index to Account for Variations in Author Contribution. *Cureus 10*(5). doi:10.7759/cureus.2629

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, *26*(5), 445–458. https://doi.org/10.1080/09645292.2018.1484426

Rajah-Kanagasabai, C. J., & Roberts, L. D. (2015). Predicting self-reported research misconduct and questionable research practices in university students using an augmented Theory of Planned Behavior. *Frontiers in psychology*, *6*, 535. https://doi.org/10.3389/fpsyg.2015.00535

Rawat, S. (2022, June 6). *There is no replication crisis in science. It's the base rate fallacy.* Big Think. https://bigthink.com/hard-science/the-replication-crisis-is-overstated/?utm_source=chatgpt.com

Resnik D. B. (2011). Scientific research and the public trust. *Science and engineering ethics*, *17*(3), 399–409. https://doi.org/10.1007/s11948-010-9210-x

Romanovsky A. A. (2012). Revised h index for biomedical research. *Cell cycle (Georgetown, Tex.)*, *11*(22), 4118–4121. https://doi.org/10.4161/cc.22179

Rooprai, P., Islam, N., Salameh, J.P., Ebrahimzadeh, S., Kazi, A., Frank, R., Ramsay, T., Mathur, M.B., Absi, M., Khalil, A., Kazi, S., Dawit, H., Lam, E., Fabiano, N. and McInnes, M.D. (2023) 'Is there evidence of p-hacking in imaging research?', *Canadian Association of Radiologists Journal*, 74(3), pp. 497–507. Available at: https://doi.org/10.1177/08465371221139418 (Accessed: 3 June 2026).

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rubin, M. (2022) 'The costs of HARKing', *The British Journal for the Philosophy of Science*, 73(2), pp. 535–560. Available at: https://doi.org/10.1093/bjps/axz050 (Accessed: 3 June 2026).

Rubin, M. (2023) 'Questionable metascience practices', *Journal of Trial & Error*, 4(1), pp. 5–20. Available at: https://doi.org/10.36850/mr4 (Accessed: 3 June 2026).

Rubin, M. (2025) *A brief review of research that questions the impact of questionable research practices*. PsyArXiv. Available at: https://doi.org/10.31234/osf.io/ah9wb_v3 (Accessed: 3 June 2026).

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In Defense of the Questionable: Defining the Basis of Research Scientists' Engagement in Questionable Research Practices. *Journal of empirical research on human research ethics : JERHRE*, *13*(1), 101–110. https://doi.org/10.1177/1556264617743834

Satalkar, P., & Shaw, D. (2019). How do researchers acquire and develop notions of research integrity? A qualitative study among biomedical researchers in Switzerland. *BMC medical ethics*, *20*(1), 72. https://doi.org/10.1186/s12910-019-0410-x

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology, 13*(2), 90–100. https://doi.org/10.1037/a0015108

Schneider, J. W., Allum, N., Andersen, J. P., Petersen, M. B., Madsen, E. B., Mejlgaard, N., & Zachariae, R. (2024). Is something rotten in the state of Denmark? Cross-national evidence for widespread involvement but not systematic use of questionable research practices across all fields of research. *PloS one*, *19*(8), e0304342. https://doi.org/10.1371/journal.pone.0304342

Sherif, M. (1936). *The psychology of social norms.* Harper.

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. https://doi.org/10.1177/0956797611417632

Smyth, M. L., & Davis, J. R. (2004). Perceptions of Dishonesty among Two-year College Students: Academic versus Business Situations. *Journal of Business Ethics*, *51*(1), 63-73. https://doi.org/10.1023/B:BUSI.0000032347.79241.3c

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American Statistical Association, 54,* 30–34. https://doi.org/10.2307/2282137

Sterling, T. D., Rosenbaum, W. L., & Weinkam, J. J. (1995). Publication Decisions Revisited: The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa. *The American Statistician*, *49*(1), 108–112. https://doi.org/10.1080/00031305.1995.10476125

Stone, T. H., Jawahar, I. M., & Kisamore, J. L. (2009). Using the theory of planned behavior and cheating justifications to predict academic misconduct. *The Career Development International, 14*(3), 221–241. https://doi.org/10.1108/13620430910966415

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651

The Concordat Working Group. (2016). *Concordat on open research data*. https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf

Tijdink, J. K., Bouter, L. M., Veldkamp, C. L., van de Ven, P. M., Wicherts, J. M., & Smulders, Y. M. (2016). Personality Traits Are Associated with Research Misbehavior in Dutch Scientists: A Cross-Sectional Study. *PloS one*, *11*(9), e0163251. https://doi.org/10.1371/journal.pone.0163251

Tijdink, J. K., Vergouwen, A. C., & Smulders, Y. M. (2013). Publication pressure and burn out among Dutch medical professors: a nationwide survey. *PloS one*, *8*(9), e73381. https://doi.org/10.1371/journal.pone.0073381

Ulrich, R. and Miller, J. (2020) 'Questionable research practices may have little effect on replicability', *eLife*, 9, e58237. Available at: https://doi.org/10.7554/eLife.58237 (Accessed: 3 June 2026).

University of Oxford. (2023, 18 September). *Research integrity online training.* https://researchsupport.admin.ox.ac.uk/article/research-integrity-online-training

Van Noorden R. (2023). More than 10,000 research papers were retracted in 2023 - a new record. *Nature*, *624*(7992), 479–481. https://doi.org/10.1038/d41586-023-03974-8

Vancouver, J.B. (2018) 'In defense of HARKing', *Industrial and Organizational Psychology*, 11(1), pp. 73–80. Available at: https://doi.org/10.1017/iop.2017.89 (Accessed: 3 June 2026).

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. (2011). Towards a new crown indicator: an empirical analysis. *Scientometrics*, *87*(3), 467–481. https://doi.org/10.1007/s11192-011-0354-5

Ware, J. J., & Munafò, M. R. (2015). Significance chasing in research practice: causes, consequences and possible solutions. *Addiction (Abingdon, England)*, *110*(1), 4–8. https://doi.org/10.1111/add.12673

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid *p*-hacking. *Frontiers in Psychology, 7,* Article 1832. https://doi.org/10.3389/fpsyg.2016.01832

Woznyj, H. M., Grenier, K., Ross, R., Banks, G. C., & Rogelberg, S. G. (2018). Results-blind review: A masked crusader for science. *European Journal of Work and Organizational Psychology, 27*(5), 561–576. https://doi.org/10.1080/1359432X.2018.1496081

Wu, Q. (2009). The w-index: A measure to assess scientific impact by focusing on widely cited papers. *Journal of the American Society for Information Science and Technology*. *61*(3), 609–614. https://doi.org/10.1002/asi.21276

Xie, Y., Wang, K., & Kong, Y. (2021). Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis. *Science and engineering ethics*, *27*(4), 41. https://doi.org/10.1007/s11948-021-00314-9

Zeng, L., Yu, X., Yu, T., Xiao, J., & Huang, Y. (2019). Interventions for smoking cessation in people diagnosed with lung cancer. *The Cochrane database of systematic reviews*, *6*(6), CD011751. https://doi.org/10.1002/14651858.CD011751.pub3

Chapter 2

Investigating the links between questionable research practices, scientific norms and organisational culture

Paper A was published in *Research Integrity and Peer Review* in October 2024

Abstract

Background:

This study investigates the correlates of self-reported engagement in questionable research practices (QRPs), focusing on both individual-level factors (such as scholarly field, commitment to scientific norms, gender, contract type, and career stage) and institution-level factors (including industry type, researchers' perceptions of their research culture, and awareness of institutional policies on research integrity).

Methods:

Using a multi-level modelling approach, we analyse data from an international survey of researchers working across disciplinary fields to estimate the effect of these factors on self-reported QRP engagement.

Results:

Our findings indicate that contract type, career stage, academic field, adherence to scientific norms and gender significantly predict self-reported QRP engagement. At the institution level, factors such as being outside of a collegial culture and experiencing harmful publication pressure, and the presence of safeguards against integrity breaches have small associations. Only a minimal amount of variance in self-reported QRP engagement is attributable to differences between institutions and countries.

Conclusions:

We discuss the implications of these findings for developing effective interventions to reduce QRPs, highlighting the importance of addressing both individual and institutional factors in efforts to foster research integrity.

Key Words: Research Integrity, Questionable Research Practices, Scientific Norms

Introduction

Questionable Research Practices (QRPs) refer to suboptimal research practices that exist in an area of ethical ambiguity between sound scientific conduct and outright scientific misconduct (i.e., fabrication, falsification, and plagiarism; John et al., 2012; Simmons et al., 2011; Wicherts et al., 2016; Bakker et al., 2012; LeBel et al., 2013). Given their inherent ethical ambiguity, these practices may be seen as warranted or defensible in certain contexts and some of the time (John et al., 2012). These practices include hypothesising after results are known (termed 'Harking'; Kerr, 1998), including authors on publications who have not contributed sufficiently to warrant authorship and selectively reporting study outcomes ('cherry picking'). While these practices are perhaps insignificant in isolation, their cumulative impact over time can adversely affect science by undermining the reliability and validity of scientific knowledge. For example, the various methodological decisions that are inherent to the research process, known as 'researcher degrees of freedom,' have been found to elevate Type I error rates (Simmons et al., 2011, pp.1359). QRPs lead to a skewed scientific literature that inevitably undermines efficiency and impedes the scientific process. That is, prolonging support for empirically untenable theories that should be revised or discarded. Downstream, these practices at the very least contribute to wasted resources and at most lead to the canonisation of erroneous scientific claims (Jussim et al., 2019), the compromising of science-informed policy and interventions, and an erosion of public trust in science (Jussim et al., 2019).[6]

---

[6] The relationship between QRPs and replication is contested. Some research suggests that QRPs have limited effects on replication outcomes (Ulrich & Miller, 2020), operate mainly in specific contexts (Rubin, 2023; Rubin, 2025), or may even increase apparent replication rates (Francis, 2012). Thus, practices such as p-hacking, HARKING, and publication bias may be less central to the replication crisis than often assumed (Rubin, 2025), and both their harmfulness and the significance of low replication rates have been debated.

The author therefore recognises that not all QRPs are unethical or harmful, and that some may be justifiable in context. Nevertheless, QRPs may still have broader consequences, including false literatures, wasted resources, theoretical distortion, poor educational signals, QRPs are of particular concern given the frequency at which researchers engage in them. Martinson et al. (Martinson et al., 2005) found that 33% of respondents in their survey admitted to having engaged in at least one QRP or research misconduct within the last three years, while John and colleagues (John et al., 2012), using an anonymous elicitation survey format combined with incentives for truth telling, identified prevalence estimates up to 100% for some QRPs amongst psychologists. Like Martinson and colleagues (2005), an early meta-analytic study by Fanelli (2009) found that approximately 34% of respondents admitted to engaging in at least one QRP at some point during their academic career. However, a recent meta-analytic study by Xie and colleagues (2021) found a more modest estimate, with prevalence rates of around 12.5%. While there is evidence to suggest that the interpretation of survey data can sometimes lead to exaggerated prevalence estimates, as in the case with John and colleagues' (2012) study (see, Fiedler & Schwarz, 2016), the existence of prevalence rates that are far from non-zero is a matter of considerable concern.

Given the far-reaching downstream consequences and frequency of occurrence, a growing number of research publications, which have informed reform initiatives and interventional and educational strategies, have sought to elucidate and capture the factors that motivate self-reported engagement in suboptimal research practices. Proposed determinants of QRPs can be categorised across three distinct, but interrelated, strata: individual-level factors that are intrinsic to researchers, organisational-level factors pertaining to research institutions, and systemic-level factors inherent within the broader research ecosystem. The relationship between QRPs and determinants at all levels remains largely unclear, with scant empirical literature investigating the antecedents of QRPs.

---

and reduced credibility. Accordingly, this discussion treats QRPs as practices that may undermine research integrity beyond their effects on replication.

Individual Factors

At the individual-level, cognitive biases when navigating researcher degrees of freedom (MacCoun, 1998; Nuzzo, 2015; John et al., 2012) and competency shortfalls in research ethics (Johnsson et al., 2014; Resnik, 2011), research methodology (Begley & Ioannidis, 2015; Munafò et al., 2017) and data analysis (MacCoun, 2022) have been postulated as potential explanatory factors for engaging in QRPs. The personality traits of agreeableness and conscientiousness have been found to be inversely associated with self-reported prevalence of QRPs in survey research (Schneider et al., 2024; Giluk & Postlethwaite, 2015), while the trait of 'Machiavellianism' has been identified as a potential risk factor for research misbehaviour (Tijdink et al., 2016). Janke and colleagues (2019) found that among a sample of psychologists stronger appearance approach goals (i.e., striving for skill demonstration) and lower learning approach goals (i.e., striving for skill development) predicted greater engagement in self-reported QRPs. Moreover, Sacco and colleagues (2018) found that researchers judged QRPs to be ethically defensible and were more willing to engage in them where the practice was perceived to be professionally appropriate.

A researcher's commitment to normative ideals of science has the potential to subvert other motivating factors. Robert Merton (1979) argued that scientists adhere to a set of informal cultural norms, including universalism (evaluating claims based on pre-established impersonal criteria, rather than nationality, ethnicity, gender or professional affiliation), communalism (scientific findings belong to the entire scientific community), disinterestedness (scientists' work should not be influenced by personal or monetary bias, and they should work for the benefit of science), and organised scepticism (assessment of research should be based on impersonal critical scrutiny). Although recent investigations into the extent to which scientists subscribe to these norms are scant, where surveys have been carried out, primarily in specific disciplines, evidence has shown that adherence, although not universal, is widespread

(Anderson et al., 2010; Bray & von Storch, 2017; Macfarlane & Cheng, 2008; Kim & Kim, 2018). Feasibly, a commitment to these norms may mitigate the influence of competing influences, such as organisational and systemic pressures to cut corners to remain competitive. However, empirical inquiry investigating how commitment to the ideals of science is associated with behaviour is noticeably absent, except for a recent survey of researchers in The Netherlands that found that commitment to the normative ideals of science was one of the strongest predictors for not engaging in QRPs (Kim & Kim, 2018).

The relationship between self-reported QRP engagement and other individual-level factors, such as gender, career stage and years since obtaining a PhD are even less clear. Preliminary evidence from a sample of researchers in The Netherlands found that women are less likely to engage in QRPs (Gopalakrishna et al., 2022), while a more recent study by Schneider et al. (2024) found that male Danish respondents reported slightly higher QRP prevalence rates compared to female respondents. This effect was less clear among a non-Danish sample. In the same study, Schneider and colleagues found that academic experience (years since obtaining a PhD) was associated with less self-reported QRP engagement. Relatedly, Gopalakrishna et al. (2022) found that PhD candidates and junior researchers are more likely to engage in at least one QRP frequently compared to postdocs and assistant professors. However, no statistically significant relationship between academic rank and average self-reported engagement in QRPs was found. Research showing disciplinary field differences in self-reported QRP engagement additionally present conflicting findings (e.g. Xie et al., 2021; Schneider et al., 2024).

Organisational and Systemic Factors

At the system level, it is widely believed that the misalignment between incentive structures and the principles of scientific integrity play a significant role in influencing the self-reported engagement in QRPs (e.g., Martinson et al., 2005; Bouter, 2015; Edwards & Roy, 2017). The favouring of the aesthetic quality of results (e.g., novelty and statistical significance) by journal editors over the reproducibility and robustness of scientific findings arguably leads to a culture that prioritises eye-catching outcomes at the expense of methodological soundness. This is demonstrated by the conspicuous absence of null findings in the scientific literature. Chavalarias and colleagues (2016) found that statistically significant results (i.e., p-value < 0.05) account for approximately 96% of published abstracts in MEDLINE and full-text articles in PubMed Central. On this basis, there have been calls for journals to increase publication of negative results (Ware & Munafò, 2015; Nosek et al., 2012) and confirmatory finding (Higginson & Munafò, 2016), as well as improved systems and platforms to facilitate reproducibility (Samota & Davey, 2021). Moreover, registered reports, which involve journal editors agreeing to publish unknown results based on a proposed design and methodological rigour, are increasingly being utilised (e.g., in Cortex and Comprehensive Results in Social Psychology) to address the bias towards statistically significant findings.

At the institutional-level, simple metrics, such as frequency of publication and journal impact factors, are used when assessing for hiring, promotion, tenure, and permanency. These institutional metrics combined with the systemic bias that favours the aesthetic quality of research, creates a competitive and high-pressure culture, where researchers are incentivised to cut corners to remain competitive. Several studies have found self-reported publication pressure to be positively associated with self-reported use of QRPs (Gopalakrishna et al., 2022; Maggio et al., 2019; Schneider et al., 2024). Research in the Netherlands has shown that this pressure is felt by researchers operating across fields, with approximately 54% of surveyed professors feeling that the pressure to publish is excessive (Tijdink et al., 2013). On this basis, there have been calls for broad institutional revisions in how researchers are hired, evaluated, and promoted (Kretser et al., 2019; Moher, 2018), with a reduced emphasis on journal impact factors and other commonly used metrics (Diaz-Martinez, 2013). Similarly, Principles 4 and 5 of the Hong Kong Manifesto endorsed by the 6th World Conference on Research Integrity call for reforms in the way researchers are assessed and valued, focusing on valuing a broader range of researcher contributions, including replication attempts, peer review, mentoring and knowledge transfer (6^{th} World Conference on Research Integrity, 2019). Empirically, local culture has been found to be associated with QRP prevalence (Schneider et al., 2024), whereby research cultures that promote quality and rigorous research, and reward integrity, appear to counter systemic use of QRPs to some degree.

These organisational and system level factors may interact with individual-level motivations, such as the personal-economic and career-specific motivations of researchers. For example, institutions that pressure scientists to increase their research output (e.g., using publication metrics) are most likely especially influential when scientists are motivated by job permanency, promotions, pay rises and social approbation. That is, the susceptibility to negative impacts (e.g., systemic influences) on behaviours of researchers may not be uniformly distributed across individuals (Schneider et al., 2024).

The Current Study

In this study, we seek to explore the relative influence of selected individual-level, organisational and systemic level factors on self-reported engagement in QRPs, using data from an international survey of researchers operating across scholarly fields. The factors suggested as explanations for self-reported engagement in QRPs in the extant research literature are grounded in limited and disjointed evidence. We aim to bring a more coherent and broader perspective by exploiting a comprehensive source of data not previously available. Additionally, the relative importance of individual versus organisational factors in explaining QRPs is an important, but under-explored, area of inquiry. The relative influence attributable to individual-level and organisation-level factors is important for the development of efficacious interventional strategies that appropriately target the primary motivators of self-reported QRP engagement, thus improving the efficiency, quality, and credibility of scientific research. On the one hand, if a significant proportion of variance in researcher self-reported QRP engagement is attributable to organisational-level differences, there is an impetus for further examination of institutionally specific factors and institutionally tailored solutions and policy revisions within these settings. On the other, if systematic variation in QRPs across institutions is minimal, this would suggest that individual-level factors (e.g., nurturing ideals of sound science and competencies), as well as systemic factors (e.g., eliminating systemic pressures and hyper-competition), should be the primary targets of intervention.

The International Survey on Research Integrity (IRIS), a large-N cross-sectional survey, administered to researchers in the US and Europe, presents an ideal opportunity to assess the impact of some widely postulated individual and organisational level factors in motivating self-reported QRP engagement. IRIS includes individual-level items allowing us to explore the impact of employment contract type, career stage, disciplinary field, sex, and commitment to the ideals of science on self-reported QRP engagement. It additionally includes items that measure organisational level features, including workplace type, existence of integrity training and procedures for tackling integrity breaches, whether institutions have a written statement on research integrity and perceptions of collegiality and pressure in working environment. Thus, the survey data permits us to explore the relative impact of organisational-level characteristics on self-reported QRP engagement, at least as perceived by the respondent. The ability to disaggregate research institutions additionally allows us to explore the extent to which variance in self-reported QRP engagement can be attributed to organisational and individual level differences. Moreover, while existing research typically examines a narrow set of disciplinary fields, the IRIS includes respondents representing researchers from across diverse fields.

The cross-national nature of the survey allows us to go beyond individual and organisation and examine country level differences in self-reported QRP engagement. Xie, Wang and Kong (2021) found that research misconduct and QRPs appear to be higher in low- and middle-income countries (LMICs, e.g., Nigeria and India), relative to high-income countries (HICs, e.g., the UK and USA), resulting from lack of provision of standards of responsible research conduct and lack of organisation. However, there are also distinct standards, policies and cultures that could differentiate amongst HICs. For example, Allum et al, using the IRIS data, found differences between Europe and the US on a range of research integrity indicators, including QRPs. We go beyond their analysis to consider how important is the country context after accounting for individual characteristics – which may be distributed quite differently across HICs. Diverse educational systems with differing emphases placed on ethical practices, as well as differing regulatory frameworks for detecting and deterring self-reported QRP engagement are among those unobserved factors that may lead to substantial country level variance. Our analysis permits us to explore these country differences, predominantly within Europe, alongside the other explanatory variables already described.

Data and Methods

The data come from a cross-sectional international survey of researchers, administered in 2021 as part of the European Commission-funded ‘Standard Operating Procedures for Research Integrity’ (SOPs4RI) project. The online survey utilised a systematic, stratified probability sample of over 60,000 researchers from across scholarly fields. The sample comprised authors of research articles indexed in Clarivate’s Web of Science citation database. The survey protocol was pre-registered with the Centre for Open Science in February 2019 (available, along with reproducibility materials at Allum et al., 2024). Further details about IRIS are contained in Allum et al. (2023) Our analytical sample is comprised of 39,699 researchers. Of these, 57.6% were male. 40.8% primarily research in the natural sciences, 14.4% in the medical sciences, 30.8% in the social sciences and 14% in the humanities. 4.5% are employed in the United States of America, 86.7% in Europe and 8.7% elsewhere.

Variables

Questionable Research Practices. A total of eight behaviours pertaining to various aspects of the research process (e.g., publication, PhD supervision) that are widely considered to be questionable and undermining of the trustworthiness of scientific findings were generated for the purposes of the survey. Descriptions of these questionable research practices (QRPs) can be found in Table 2.1.

Each QRP was accompanied with a question asking: 'thinking about your research carried out for your publications over the last three years, how often has the following occurred?', with response categories 'often', 'sometimes', 'rarely', 'never', 'does not apply'. The three-year timeframe was used to limit recall bias and to be able to state prevalence in more precise terms. Moreover, Fiedler and Schwarz (2016) argue that traditional QRP surveys frequently produce inflated and misleading prevalence estimates by measuring lifetime binary exposure (i.e., asking whether the respondents has *ever* engaged in the practice). These surveys conflate the proportion of individual scientists who have made a single error with the true systemic frequency of the behaviour across all scientific projects. On that basis, we utilise a three-year time bound that evaluates the *extent* of self-reported engagement, rather than a lifetime binary metric. Accordingly, the data and models presented herein should be interpreted strictly as evidence of associations and behavioural patterns within the responding sample. The findings are intended to illuminate the modifiable features of the research environment and incentive structures that predict these behaviours, rather than to serve as a census of absolute prevalence or a definitive index of systemic replicability.

The eight individual QRP item scores were aggregated for each respondent, resulting in a total score ranging from 0 to 32. This sum was then divided by the count of items that were both responded to and deemed relevant by respondent (i.e., where 'does not apply' was not selected). This scale was standardised to z-scores, ensuring each score is mean-centered and scaled by the standard deviation, with higher scores indicating greater average self-reported engagement. Operationalising 'QRP engagement' in this way is consistent with previous research that has attempted to model the relationship between various explanatory factors and QRPs (e.g., Gopalakrishna et al., 2022). Figure 2.1 displays the weighted distribution of our standardised mean self-reported QRP engagement (mean = -0.04, median = -0.18, standard deviation = 0.92).

Table 2.1. Questionable Research Practice items from International Survey on Research Integrity

	Questionable Research Practice (QRP) Question Wording
QRP 1	Wilfully failing to cite relevant publications that contradict your own beliefs, theories, hypotheses, methods or findings.
QRP 2	When reviewing a manuscript, not investing the effort necessary to conduct a thorough review.
QRP 3	Choosing not to report your findings if they could weaken or contradict your theories of hypotheses.

QRP 4	Deliberately using another researcher's unpublished idea without giving credit. For example, publishing an idea voiced by a colleague at an informal meeting without giving them credit.
QRP 5	In a publication, failing to disclose relevant personal, financial, political or intellectual conflicts of interests.
QRP 6	Including authors on a paper who had not contributed sufficiently to the work to merit authorship.
QRP 7	Inadequately supervising or mentoring junior co-workers.
QRP 8	Carrying out research without getting the required ethical approval.

Respondent-level Explanatory Factors. A composite variable indicating scientists' commitments to the normative ideals of science was computed based on five items assessing researchers' values, each based on Merton's delineation of the normative ideals of science (i.e., universalism, disinterestedness, organised scepticism, and communalism). These items were adapted from Anderson (Anderson, 2000) and Anderson et al. (2010). They described a set of behaviours, and respondents indicated whether they personally feel that these behaviours reflect how researchers should behave, with response categories anchored at 'yes, usually should' (coded 1) to 'no, never should' (coded 5). Descriptions of these normative ideals can be found in Table 2.2. Items 3 and 4 were reverse-coded. For each respondent, an average score was computed based on the five items. This composite variable was standardised to mean zero and standard deviation one. Higher scores indicate a greater commitment to the normative ideals of science. Figure 2.2 displays the weighted distribution of our standardised composite variable, which represents the average adherence to the normative ideals of science (mean = 0.06, median = 0.12, standard deviation = 0.88).

Figure 2.1 Weighted distribution of standardised mean self-reported QRP engagement for 39,699 respondents

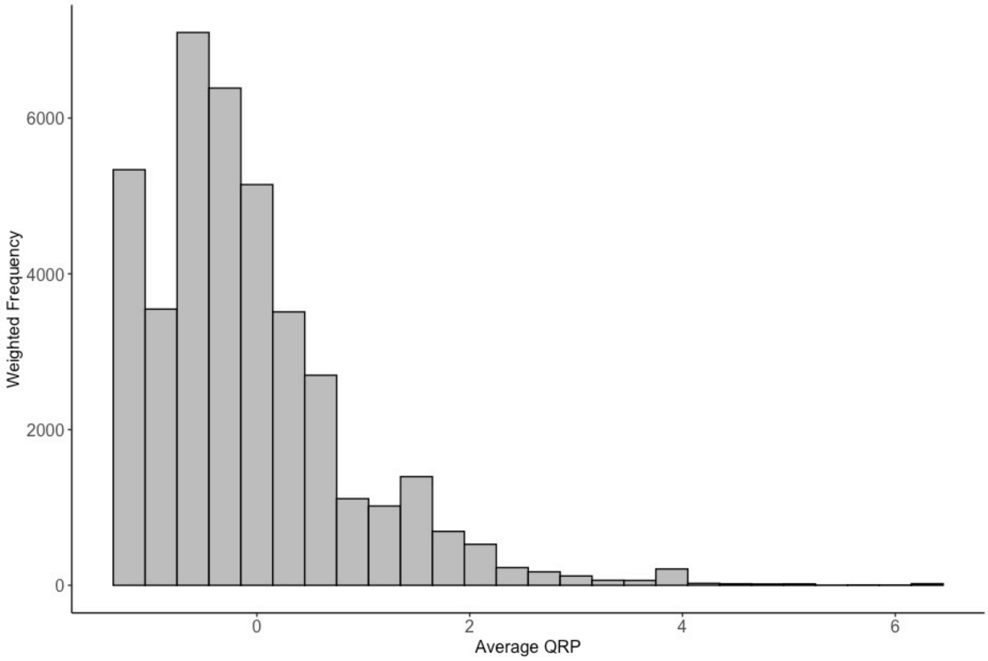
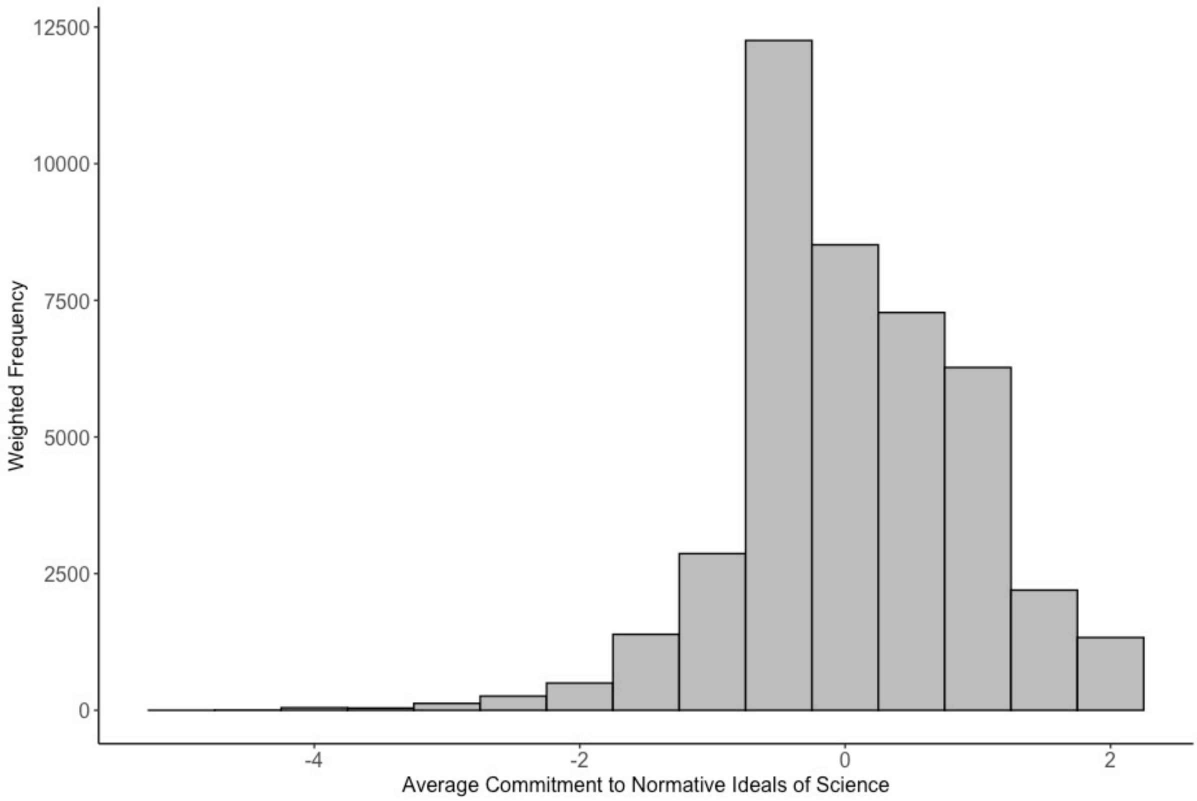


Table 2.2. Normative ideals of science items from International Survey on Research Integrity

	Mertonian Norm	Normative Ideal of Science Question Wording
Item 1	Communalism	"Researchers should openly share new findings with colleagues"
Item 2	Organised Scepticism	"Researchers should consider all new evidence"
Item 3	Disinterestedness 1	"Intellectual work should be influenced by personal beliefs and values"
Item 4	Disinterestedness 2	"Researchers should change their research interests to access research funding"
Item 5	Universalism	"Researchers should always publish findings that are scientifically sound"

Each respondent reported their main disciplinary field and was assigned to one of four broad categories: social sciences, humanities, medical sciences (including biomedicine), and natural sciences (including technical sciences). This variable was recoded into three binary variables representing each scientific field, with 'natural sciences' omitted, as the reference category. Respondents additionally provided the type of employment contract that they are on from a pre-specified list: permanent contract, temporary contract, or no contract. We recoded this variable into two binary variables with permanent contract omitted as reference category. Respondents indicated the stage of their career, selecting from either early-career, mid-career, late-career or retired. This variable was also recoded into three dummy variables, with the reference category early-career. Each respondent reported their sex, with response categories male, female and prefer not to say. Those stating prefer not to say (N = 1109) were removed from the analysis. This was recoded as a dummy variable with male omitted as the reference category.

Figure 2.2. Weighted distribution of our standardised mean commitment to normative ideals of science for 39,699 respondents



Organisation Level Explanatory Factors. Each participant was asked to identify their workplace type from six sectors: academia, industry, not-for-profit research institute, government research centre, healthcare setting, or other. We recoded this variable into five indicator variables, using academia as the reference category. Respondents were also presented with a series of descriptions characterising what could be regarded as a functionally optimal working environment and were asked the extent to which the descriptions resemble their own working environment. The items selected for use in our analysis were the presence of adequate integrity training (i.e., 'training in research integrity is provided to all researchers, at all career stages, by qualified trainers'), provision for handling integrity breaches (i.e., 'researchers can consult a qualified person in confidence with any research integrity concerns. Breaches are detected and sanctioned in a fair and standardized way, protecting both whistleblowers and those accused of misconduct') and a positive working culture (i.e., 'collegial, and without harmful publication pressure, detrimental power imbalances or conflict'). These were measured on a five-point scale, anchored at 'resembles my environment very closely' (coded 1) and 'resembles my environment not at all closely' (coded 5). We reverse coded this variable so that higher scores indicated that the description is more reflective of their working environment. The variables were mean-centred and standardised to z scores. Additionally, respondents were asked to state whether their institution has a written statement on research integrity, with three answer categories: 'yes', 'no' and 'I don't know'. We recoded this variable into two dummy variables, with 'yes' as the reference category.

Contextual Variables. Integral to the present research is the degree to which variability in self-reported QRP engagement is associated with differences between organisations and countries. We use the email addresses of survey respondents to create a variable indexing organisations, based on their domain name (e.g. @harvard.edu). Generic email addresses were identified using a non-exhaustive list of the most used general email addresses (e.g. Gmail, Hotmail). We deleted cases where the domain name was not associated with a research-producing organisation. In total, 9,304 non-institutional email addresses were removed, leaving our analytical sample containing a total of 7,666 unique organisations. Figure S2.1 in the supplementary material presents the frequency distribution and descriptive statistics for organisation. The second contextual variable, country, was based on the self-reported country of employment. We focused on the 'country of employment' as it reflects the researcher's location and, consequently, their organisational setting. The final analytical sample consisted of respondents employed across 34 countries. These were predominantly in Europe.

Interaction terms. To assess the extent to which a researcher's scientific values may act as a safeguard against workplace pressures and non-collegial working environments, thus reducing self-reported engagement in QRPs, we specify a cross-level interaction, interacting commitment to the normative ideals of science with working environment. We additionally test whether a researcher's working environment has more of an influence on research conduct for those early in their career and those on non-permanent contracts, who potentially have more to gain from engaging in suboptimal practices. On this basis, we create several interaction terms, interacting the working environment on our career stage and employment contract indicator variables. Moreover, we expect the influence of a less collegial and more competitive working environment on self-reported engagement in QRPs to be reduced where integrity training is sufficient, there are procedures (incl. firmer penalties) for handling contraventions of good research practice, and where the researcher is aware of the institutions commitment to ensuring research integrity (i.e., by way of awareness of a written statement on research integrity). Therefore, we specify several interaction terms, interacting working environment with a) existence of integrity training, b) procedures for handling integrity breaches, c) there being no written statement on research integrity and d) lack of awareness of whether there is a written statement on research integrity.

The presence of integrity training, along with suitable procedures and personnel for addressing integrity violations, could potentially support researchers in adhering to Mertonian scientific values. As a result, we define interaction terms for the commitment to these ideals of science in relation to a) integrity training, and b) handling of integrity breaches. Finally, the relationship between having an awareness of a written statement on research integrity and self-reported QRP engagement will plausibly be stronger for those who are more strongly committed to the normative ideals of science. Conversely, we expect the presence of a research integrity statement to matter less, and be less influential on research conduct, for those who do not adhere to the normative ideals of science. On this basis, we generate two interaction terms, interacting commitment to scientific ideals of science on a) lack of awareness as to whether their institution has written statement on research integrity, and b) an awareness that their institution does not have a written statement on research integrity.

Analysis Strategy

To account for the non-independence of responses to the QRP items within organisations and countries, and in recognising the hierarchical nature of the data (researchers nested within organisations and countries), we use a multilevel regression approach, implemented using the lme4 package in R (version 2022.12.0+353). Adopting a multilevel modelling approach means that we can partition the variability of self-reported QRP engagement across individuals, organisations, and countries, allowing us to examine the relative importance of these three components.

Our modelling strategy is as follows. In Model 1 we specify a random intercept model without any predictors, allowing us to assess the variability in QRP reporting at each of our three levels. In Model 2 we again specify a random intercept model, regressing standardised mean self-reported QRP engagement on individual-level predictors (incl. contract type, career stage, disciplinary field, and sex). In Model 3 we include our individual-level variable of primary interest, commitment to the normative ideals of science. In Model 4 we include organisation-level predictors (incl. workplace type, integrity breaches, integrity training and RI statement awareness), allowing us to see whether variability in self-reported QRP engagement between organisations is explained by these specific organisation-level attributes.

In Model 5 we add our final variable of interest, working environment, into a random-slope model. Including random slopes permits us to see how the relationship between QRPs and a) normative ideals of science, and b) working environment, varies across organisations. Finally, in Model 6 we include several single-level and cross-level interaction terms, allowing us to see whether the relationship between a) commitment to the normative ideals of science, and b) working environment, is moderated by career stage, contract type, RI statement awareness, integrity training and integrity breaches. All models are multiple-linear regression models using restricted maximum likelihood (REML). We have made the full R code available on the Open Science Framework (OSF; Brooker & Allum, 2024).

Results

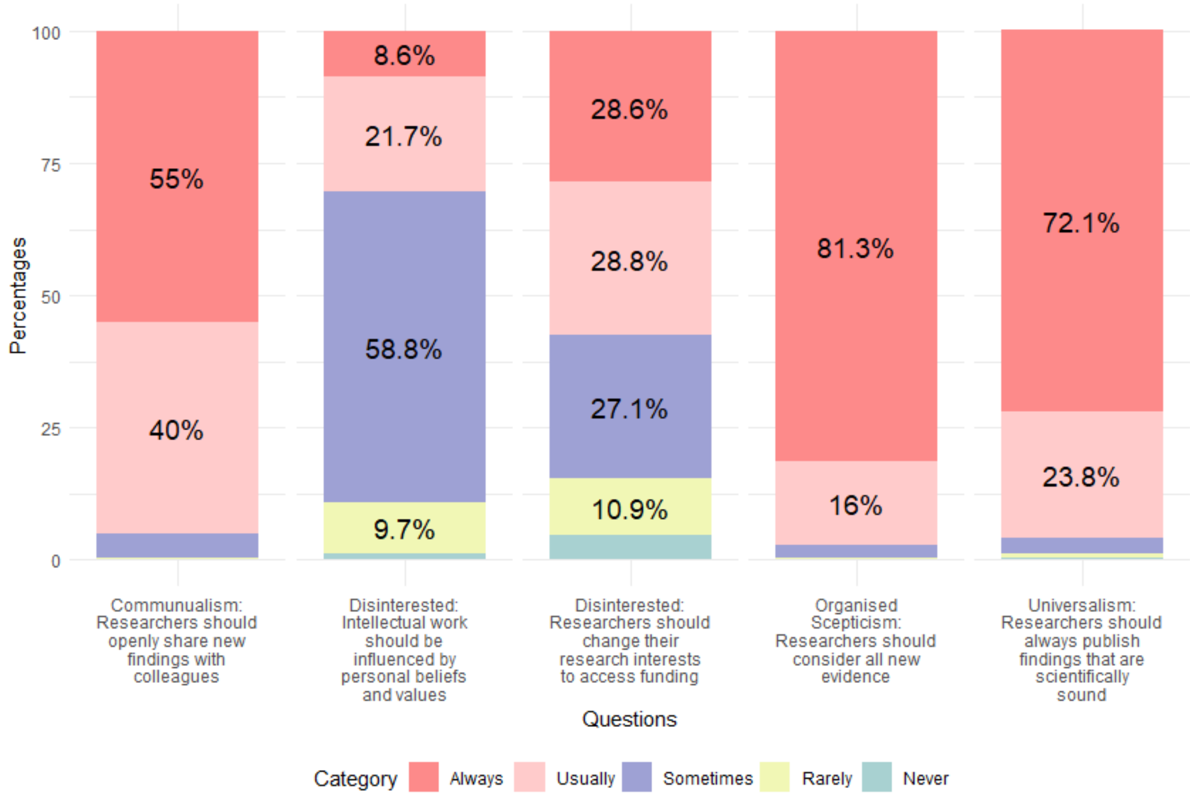
Commitment to the Normative Ideals of Science

Figure 2.3 shows the weighted percentage of respondents who support a specific practice that undermines or exemplifies the normative ideal of science, indicative of their level of commitment to these norms. A small majority of researchers are demonstrably committed to the normative ideals of communalism, organised scepticism, and universalism, with most respondents agreeing that researchers should always or usually engage in behaviours that are indicative of support to these ideals. Support for disinterestedness is less clear, with a small majority stating that researchers should change their research interests to access funding, and a sizeable majority stating that at least sometimes intellectual work should be influenced by personal beliefs and values. In line with previous research in this area, support is not universal, with a small but significant minority failing to demonstrate commitment to the normative ideals of science.

Multi-level Modelling

Table 2.3 shows the model fixed effect estimates for mean frequency of self-reported engagement in questionable research practices (QRPs) for Models 2 to 6. Model's 2 and 3 show the individual-level fixed effects and random components, while models 4 and 5 introduce our organisational-level predictors. Model 6 includes our interaction terms to test for moderating effects. Model 1 is the intercept only model. The main interest here is in the variance components. The percentage of variance due to country and organisation is extremely small – 0.95% of the variability in average self-reported engagement in QRPs is situated at the organisational level and 1.69% of the variability at the country-level. The overwhelming majority of variation in self-reported QRP engagement is between individual researchers regardless of where they are.

Figure 2.3. Commitment to normative ideals of science: weighted percentages



Model 2 introduces individual-level predictors. Researchers in the social sciences (β = -0.09) and humanities (β = -0.29) engaged in fewer QRPs on average in comparison to those in the natural sciences. In comparing the medical sciences to the natural sciences, there was a smaller yet statistically significant difference in the self-reported engagement with QRPs (β = 0.03). Individuals in the middle of their careers (β = -0.05), those later in their careers (β = -0.12) and those who are retired (β = -0.19) reported engaging in fewer QRPs on average in comparison to those early in their careers. A significant difference in average number of QRPs engaged in was observed between individuals without an employment contract (β = 0.09) and those on a temporary employment contract (β = 0.06), compared to those with a permanent contract. That is, those without an employment contract or those on a temporary contract engage in a greater number of QRPs on average compared to those on a permanent contract.

However, the effect size is very small, indicating that the practical impact of type of employment contract on self-reported engagement in QRPs is relatively minor. Small sex differences are present, with women being less likely to report QRPs (β = -0.02).

Of primary interest in this paper is the extent to which a researcher's commitment to the normative ideals of science is associated with the researcher's self-reported engagement in QRPs. In Model 3, our composite indicator representing researchers' commitment these normative ideals was added. As predicted, a greater commitment to the normative ideals of science has a negative association with the standardised average number of QRPs engaged in (β = -0.15), while holding other predictors constant.

Organisational-level predictors were added as covariates in Model 4. Average self-reported QRP engagement was observed to be greater among individuals working in industry (β = 0.19) and non-profit organisations (β = 0.12), government research centres (β = 0.06) and health research centres (β = 0.06), in comparison to those working in academia. There was no relationship found between researchers' perceptions of one's environment more closely resembling one that provides research integrity training and average self-reported engagement in QRPs. The presence of procedures and qualified staff to handle research integrity breaches was negatively related to average number of QRPs engaged in (β = -0.08). Having no awareness of whether one's institution has a written statement on research integrity is associated with a higher average self-reported engagement in QRPs (β = 0.02). Reporting that an individual's institution does not have a written statement of research integrity is associated with a higher average engagement in QRPs (β = 0.09). Again, the magnitude of the effect sizes is small. The data suggests that a researcher's normative commitment to the ideals of science robustly predicts the extent of self-reported engagement in QRPs. This is more pronounced than the impact of factors such as provision of research integrity training at one's institution, the existence of procedures and qualified staff to address research integrity breaches, or an awareness of a research integrity statement. Of note here is the resilience of a researcher's commitment to normative ideals against all other covariates that are introduced.

The extent to which the perceived collegiality and degree of harmful publication pressure affects self-reported engagement in QRPs was of particular interest to us. On this basis, the extent to which an environment was perceived to be collegial, without harmful publication pressure, harmful power disparities and conflict was added as a covariate in Model 5. The more collegial an individual's working environment is, the less individuals report engaging in QRPs (β = -0.10). Figure 2.4 illustrates predicted mean self-reported QRP engagement according to both commitment to scientific norms and the extent to which a researcher's working environment is perceived as collegial, our predictors of particular interest. The x-axis shows a range of standardised values for both of these predictors, with the red and blue lines representing norms and environment respectively. The y-axis represents the standardised mean self-reported QRP engagement. A notable negative relationship between self-reported QRP engagement and both of these predictors is visible. This suggests that researchers who adhere more closely to scientific norms tend to engage less in QRPs. Moreover, the extent to which a researcher's environment is perceived as collegial and free from harmful publication pressure, shows a negative association with self-reported QRP engagement. Researchers working in more supportive and less pressured environments tend to report engaging less often in QRPs.

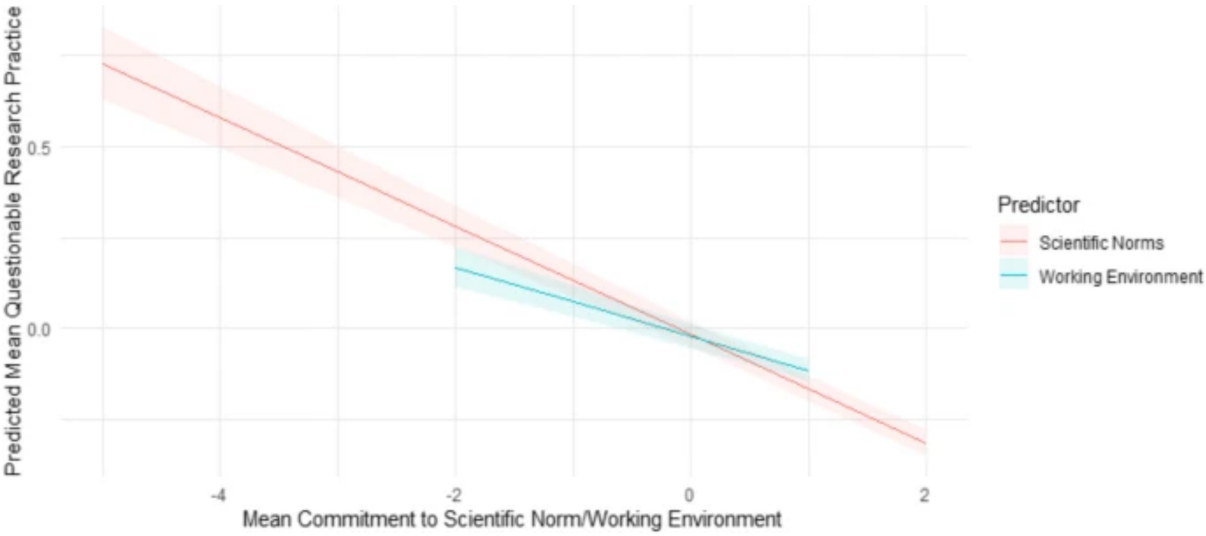
Table 2.3. Regression coefficients (95% confidence interval) of overall mean questionable research practices (QRP)

	Model 2	Model 3	Model 4	Model 5	Model 6
Level 1 Fixed Effects					
Contract Type (ref = permanent)					
no contract	0.09 (0.04, 0.14)***	0.08 (0.03, 0.13)***	0.07 (0.02, 0.12)**	0.05 (0.00, 0.10)*	0.05 (0.00, 0.09)*
temporary contract	0.06 (0.04, 0.09)***	0.06 (0.03, 0.08)**	0.07 (0.05, 0.10)***	0.06 (0.04, 0.09)***	0.06 (0.04, 0.09)***
Career Stage (ref = early career)					
mid-career	-0.05 (-0.08, -0.03)***	-0.06 (-0.08, -0.03)***	-0.07 (-0.09, -0.04)***	-0.07 (-0.09, -0.04)***	-0.07 (-0.10, -0.05)***
later-career	-0.12 (-0.15, -0.09)***	-0.12 (-0.15, -0.09)***	-0.10 (-0.13, -0.07)***	-0.09 (-0.13, -0.07)***	-0.11 (-0.13, -0.08)***
retired	-0.19 (-0.24, -0.14)***	-0.19 (-0.24, -0.14)***	-0.18 (-0.23, -0.14)***	-0.17 (-0.22, -0.13)***	-0.18 (-0.23, -0.13)***
Disciplinary Field (ref = natural sciences					
medical	0.03 (-0.00, 0.06)·	0.05 (0.02, 0.08)**	0.07 (0.04, 0.10)***	0.07 (0.04, 0.10)***	0.07 (0.04, 0.10)***
social	-0.09 (-0.12, -0.07)***	-0.11 (-0.13, -0.09)***	-0.09 (-0.11, -0.07)***	-0.10 (-0.12, -0.08)***	-0.10 (-0.12, -0.08)***
humanities	-0.29 (-0.32, -0.26)***	-0.31 (-0.33, -0.28)***	-0.29 (-0.32, -0.26)***	-0.29 (-0.32, -0.26)***	-0.39 (-0.33, -0.27)***
Sex (Ref = Male)	-0.02 (-0.04, 0.00)·	-0.03 (-0.05, -0.01)**	-0.03 (-0.05, -0.01)**	-0.04 (-0.06, -0.03)***	-0.04 (-0.06, -0.02)***
Scientific norms		-0.15 (-0.16, -0.14)***	-0.15 (-0.16, -0.14)***	-0.15 (-0.17, -0.13)***	-0.15 (-0.17, -0.14)***
Level 2 Fixed Effects					
Workplace type (ref = Academia)					
industry			0.19 (0.13, 0.26)***	0.21 (0.15, 0.27)***	0.21 (0.14, 0.27)***
non profit			0.12 (0.08, 0.16)***	0.14 (0.10, 0.18)***	0.13 (0.09, 0.17)***
government research			0.06 (0.02, 0.10)**	0.07 (0.04, 0.11)***	0.07 (0.03, 0.11)***
health research			0.06 (0.01, 0.11)*	0.07 (0.02, 0.13)**	0.07 (0.02, 0.12)**
other			0.04 (-0.03, 0.10)	0.04 (-0.03, 0.11)	0.04 (0.03, 0.10)
Integrity Breaches			-0.08 (-0.09, -0.06)***	-0.05 (-0.07, -0.04)***	-0.05 (-0.06, -0.04)***
Integrity Training			-0.02 (-0.03, -0.01)***	-0.01 (-0.02, 0.00)	-0.00 (-0.02, 0.01)
Awareness of RI statement (ref = aware)					
unaware			0.02 (0.00, 0.04)*	0.03 (0.01, 0.05)**	0.03 (0.01, 0.05)**
no statement			0.09 (0.06, 0.12)***	0.10 (0.07, 0.13)***	0.11 (0.08, 0.14)***
Working environment				-0.10 (-0.11, -0.08)***	-0.12 (-0.15, -0.10)***
Interactions					
Working environment * scientific norms					0.01 (0.00, 0.02)*
Scientific norms * integrity training					-0.02 (-0.03, -0.01)***
Working environment * integrity training					-0.01 (-0.02. 0.00)·

Working environment * integrity breaches	-0.01 (-0.02, 0.01)
Working environment * mid-career	0.05 (0.03, 0.08)***
Working environment * later-career	0.06 (0.03, 0.09)***
Working environment * retired	0.06 (0.02, 0.11)**
Working environment * temporary contract	-0.03 (-0.06, -0.01)**
Working environment * no contract	-0.03 (-0.08, 0.02)
Scientific norms * integrity breaches	-0.00 (-0.01, 0.01)
Scientific norms * no statement on RI policy	-0.04 (-0.07, -0.01)*
Scientific norms * unaware of RI policy	0.03 (0.01, 0.05)**
Working environment * unaware of RI policy	-0.01 (-0.03, 0.02)
Working environment * no statement on RI policy	-0.00 (-0.03, 0.03)

Source: International Research Integrity Survey (2021); *n* = 39,699, organisation = 7,666; country = 34; restricted maximum likelihood estimation; unweighted.
Overall mean QRP was computed as an average score based on the 8 QRP items with not applicable scores removed.

Figure 2.4. Predicted marginals plot showing effects of working environment and scientific norms on questionable research practices.



Model 6 tests for selected moderation effects. A significant interaction between an individual's perception of their working environment and those who are retired (β = 0.06), those late in their career (β = 0.06) and those in the middle of their career (β = 0.05), in comparison to those early in their career, was observed. These estimates indicate that the extent to which an environment is collegial – free from harmful publication pressure, harmful power disparities and conflict – is a more critical factor in mitigating extent of self-reported QRP engagement for early-career researchers compared to mid-career, late-career, or retired researchers. Likewise, and unsurprisingly, the harmful nature of the working environment matters more for those on temporary employment contracts (β = -0.03) in comparison to those with permanent contracts. However, and perhaps surprisingly, we found that the extent to which the working environment is perceived as collegial and free from publication pressure has no differential association with self-reported QRP engagement for those with no employment contract relative to those on a permanent employment contract.

A significant interaction was observed between the extent to which a researcher is committed to the norms of science and whether they have no awareness of whether their institution has a research integrity statement. We found that for those who are unaware of whether their institution has a written statement on research integrity, a stronger relationship exists between extent of commitment to scientific norms and extent of self-reported engagement in QRPs (β = 0.03). That is, a scientist's commitment to norms of science matters more when a researcher is unaware of whether their institution has a statement on research integrity. Conversely, when a researcher is aware that their institution does not have a research integrity statement, a researcher's commitment to the norms of science has less of an influence on the extent to which they engage in QRPs (β = - 0.04).

For researchers with a stronger commitment to the norms of science, being in a working environment that heavily emphasises research integrity training is related to a slight reduction in self-reported engagement in QRPs (β = -0.02). Research integrity training contributes to reduced self-reported engagement in QRPs in environments not otherwise characterised by publication pressures, unequal power imbalances or non-collegial relationships (β = -0.01). Surprisingly, for those working in a more collegial environment, that is free from harmful publication pressure and power imbalances, a stronger commitment to the norms of science is associated with a slight increase in self-reported engagement in QRPs (β = 0.01).

Turning to the random components, the variance explained by each of the grouping variables, organisation, and country, is very weak and relatively consistent across models. Table 2.4 shows the random effect estimates for mean frequency of self-reported engagement in QRPs for models 2 to 6. Models 1-4 are random intercept models, while models 5 and 6 are random slope models. We found that the percentage of variance explained by organisation is 0.87% and by country is 1.05%, indicating that 0.87% and 1.05% of the residual variability in self-reported QRP engagement can be attributed to differences between organisations and countries, respectively. The estimated intercept variance for organisation was 0.01 across all models, indicating that only a small amount of variability in extent of self-reported QRP engagement can be attributed to organisational differences, after accounting for other predictors in the model. Likewise, the estimated intercept variance for country was also found to be 0.01 in most models, suggesting that only a tiny amount of the variability in self-reported QRP engagement can be attributed to differences between countries. This explained variance is robust to the inclusion of group-level predictors into the model, with limited change between the null model and Model 4.

Moreover, there was minimal variability in the extent to which commitment to scientific norms and perceptions of one's working environment influenced self-reported QRP engagement. That is, the relationship remained consistent across organisations and countries. This variance is additionally robust to the inclusion of single-level and cross-level interactions in Model 6.

Table 2.4. Random effects of overall mean questionable research practices (QRP)

	Model 2	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Variance (Working Environment | Organisation)				0.01	
Variance (Scientific Norms | Organisation)				0.02	
Variance (Working Environment | Country)				0.00	
Variance (Scientific Norms | Country)				0.00	
Variance country (Intercept)	0.02	0.01	0.01	0.01	0.01
Variance Organisation (Intercept)	0.01	0.01	0.01	0.01	0.01
Variance (Residual)	0.84	0.82	0.81	0.78	0.80
% of country-level variance	1.69%	1.53%	1.05%		1.09
% of organisation-level variance	0.95%	0.78%	0.87%		0.86

Source: International Research Integrity Survey (2021); *n* = 39,699, organisation = 7,666; country = 34; restricted maximum likelihood estimation; unweighted.

Discussion

The IRIS is the largest survey of its kind. Respondents represent researchers from across all disciplinary fields and a large number of research institutions across Europe and USA. As such, the survey data allow for a more comprehensive analysis of support for the normative ideals of science and an exploration of the relationship between widely postulated individual- and organisational-level factors and self-reported QRP engagement than has thus far been possible. We find that self-reported commitment to each of the normative ideals of science is widespread, but not universal. A small majority of respondents state that researchers should sometimes change their research interests to access funding, while a substantial minority believe that intellectual work should be influenced by personal beliefs and values. This finding of non-universal, but widespread, adherence is in line with findings from other survey research (Anderson et al., 2010; Bray & von Storch, 2017; Macfarlane & Cheng, 2008; Kim & Kim, 2018). While minor divergences from these norms may at first appear benign, they can, as the present research suggests, have consequences for behaviour.

It was found that main disciplinary field was associated with self-reported QRP engagement. Respondents in the medical sciences, relative to the natural sciences, reported slightly higher levels of self-reported engagement in QRPs, however the effect sizes were small. Conversely, respondents in the social sciences and humanities reported lower levels of self-reported QRP engagement compared to those in the natural sciences. However, the effects should be interpreted cautiously. Disciplinary field is likely to capture a broad set of differences, including variation in research methods, publication norms, collaborative structures, measurement practices, and exposure to competitive or productivity-based incentives. For example, practices such as selective reporting, HARKING, excluding data, or failing to report null findings may have clearer meanings in experimental, quantitative, or hypothesis-testing research traditions than in interpretive, theoretical, historical, or qualitative research. Therefore, lower reported QRP engagement in the humanities, and to a lesser extent the social sciences, may not straightforwardly indicate greater research integrity. Rather, it may partly reflect differences in the kinds of research practices that are available, recognised, or regarded as questionable within those disciplinary contexts.

Similarly, higher reported engagement in the medical sciences relative to the natural sciences should not necessarily be interpreted as evidence of greater individual willingness to compromise research integrity. Medical research often involves high publication pressure, competitive funding environments, clinical or translational stakes, and strong incentives to produce statistically significant or practically actionable findings. It may also involve more formalised methodological conventions, making certain behaviours easier to identify as QRPs. The observed disciplinary differences may therefore reflect variation in normative standards, opportunity structures, methodological traditions, and interpretive frames, rather than simple differences in researcher misconduct.

It is therefore not possible to determine from these models whether the observed differences reflect genuine variation in willingness to engage in QRPs, differences in the opportunity to engage practices, or differences in how respondents interpret and report QRP items. However, disciplinary context is likely relevant to understanding variation in self-reported QRP engagement.

One of our most surprising findings is that only a very limited amount of variance in self-reported QRP engagement can be attributed to differences between organisations and countries. This, to us, suggests that factors beyond the organisational and national contexts play a significant role in influencing self-reported QRP engagement and, at the same time, that idiosyncratic researcher differences are much more important than the local contexts in which they work. What follows from this is that perhaps more emphasis should be placed on the broadest systemic-level factors that transcend research institutions, such as hyper-competition and publication pressures, compounded by an editorial bias for the aesthetic quality of results, as well as individual-level factors. In considering the individual-level factors captured in our analysis, we find some disciplinary-level differences, with those researching in the arts and humanities engaging in fewer QRPs on average, compared to all other disciplines. This aligns with the findings of Schneider et al. (2024). However, contrary to Schneider et al. (2024) and additionally Gopalakrishna et al. (2022), Fanelli (2009), and Pupovac and Fanelli (2015), we find that those in the medical sciences and the natural sciences engage in more QRPs on average compared those in the social sciences. Moreover, our findings conflict with the findings of Xie and colleagues (Xie et al., 2021) who find that QRPs are more common in the social sciences compared to the biomedical sciences. To the extent that researchers' commitments to responsible research practice is in part due to their training and positioning within their disciplines, it is here that the potential for change lies.

Those who are later in their career admit fewer QRPs on average, again consistent with Schneider et al. (2024). We find this on the one hand unsurprising, given that those later in their career are likely under reduced pressure to publish and obtain funding (having already secured permanency and established careers) and considering evidence to suggest that career-motivations may become less salient with age (Boumans et al., 2011). That is, the influence of systemic pressures may become less salient with academic age. On the other hand, those that did their graduate training a long time ago would have had less (or no) emphasis placed on what are now known to be detrimental practices – for instance p-hacking. So, it is somewhat surprising that the same senior researchers report fewer QRPs.

Affording additional support to the reduced pressure with seniority theory, we find that those on temporary employment contracts or without an employment contract are more likely to engage in QRPs on average, suggesting that precarity can motivate researchers to engage in questionable practices, feasibly to confer an advantage on them (i.e., increased likelihood of publication) where competition (i.e., for employment) is heightened. That is, they have 'more to gain', compared to their permanently employed colleagues.

Relatedly, the organisational-level factor, indicating the extent to which the researcher's working environment is perceived to be collegial, without harmful publication pressure, detrimental power imbalances and conflict, is associated with decreased average self-reported QRP engagement. This finding is not surprising, and again in-part supports dominant theorising that hyper-competitive working environments with harmful publication pressure motivate self-reported engagement in QRPs (Bouter, 2015). It aligns with similar research findings focusing on samples from specific disciplinary fields and countries (Latan et al., 2021; Necker, 2014; Maggio et al., 2019; Gopalakrishna et al., 2022), as well as Schneider et al.'s (2024) recent cross-national survey of researchers from across disciplinary fields and national contexts. However, the item presented in IRIS captures a more holistic view of the research environment (perhaps, indicative of local research culture), potentially diluting the specific effect of publication pressure itself. It suggests a complex interplay of factors that contribute to unethical behaviour, inclusive of, but not limited to, publication pressure. This conceptualisation additionally aligns with earlier survey research by Schneider et al. (2024), who found that local culture that promotes quality and rigorous research, and rewards integrity, potentially counter presumed systemic challenges. On this basis, our finding underscores the importance of fostering a collaborative and supportive research culture, where emphasis is placed on collaboration, peer support and scientific integrity, rather than aggressive competition and high publication output. In cultivating such a research culture, it may be possible to mitigate self-reported engagement in QRPs and enhance the quality of scientific research. Given that we found limited differences in self-reported QRP engagement between institutions, this explanatory factor likely supports the narrative of corrupting systemic structures that transcend organisational working environments. That is, the issue is likely system-wide, and not unique to particular institutions.

We find sex differences consistent with earlier research. Women appear to engage in somewhat fewer QRPs on average in comparison to men. Similarly, Gopalakrishna et al. (2022) found prevalence to be higher among males, in comparison to females. Schneider et al. (2024) found a similar relationship in their Danish sample, with men engaging in slightly higher rates of QRPs compared to women, but this finding was less certain in their international sample. More generally, our findings here align with research showing that males have lower moral standards in competitive environments (Robinson et al., 2000), are more likely to deceive where the deception benefits the deceiver (Friesen & Gangadharan, 2012; Gerlach et al., 2019) and are less risk averse (Sapienza et al., 2009).

Self-reported engagement in QRPs was higher, on average, in research institutions outside academia, including health and government research centres, private industry and non-profit organisations. Those working in 'industry' were more likely to engage in QRPs on average compared to all other sectors. This aligns with research findings that suggest researchers with private industry involvement are more likely to report engaging in research misbehaviours (Martinson et al., 2009). While the exact reasons for our findings on this are a matter of speculation, it is possible that factors other than the pursuit of knowledge and scientific inquiry (e.g., a researchers' values and goals) may influence a researcher's decision-choices. These priorities may overshadow the commitment to strict scientific method and considerations of validity and reliability, inadvertently encouraging QRPs. For example, in government and health research centres, there may be significant pressure to produce results that align with policy objectives and public health goals. Researchers working within these settings may be motivated to take shortcuts to demonstrate immediate and impactful outcomes.

Again, considerations that are antithetical to the scientific enterprise may motivate a researcher to cut corners.

Subscription to the normative ideals of science was the explanatory factor scale with one of the largest correlates of decreased self-reported QRP engagement. This suggests that a researcher's ideals and normative values may act as a safeguard against other motivating factors and be powerful drivers of researcher decision-making. This corroborates findings by Gopalakrishna et al. (2022), who found that commitment to scientific norms was the largest indicator of self-reported QRP engagement. Building on recommendations made by Gopalakrishna and colleagues (2022), as well as earlier findings from Anderson and colleagues (2010), our findings here suggest that it is crucial for institutional leaders to foster a culture that upholds and respects the normative ideals of science. Through formalised training, adequate mentoring and supervision, commitment to these ideals can be encouraged.

Conversely, our findings show that the mere provision of research integrity training does not predict decreased self-reported engagement in QRPs. However, it is crucial to note that this measure reflects the 'availability of research integrity training' at one's institution, rather than active engagement by researchers. Certainly, research showing that integrity training can influence integrity is contested and weak (Satalkar & Shaw, 2019; see, Conroy, 2020; Korfeld, 2012; Phillips et al., 2018), while early-childhood education and personality traits have been found to be better predictors of researcher behaviour than research integrity training (Satalkar & Shaw, 2019). A more reasonable conceptualisation is that the availability of training may signify an organisational culture that values research integrity. However, as we find here, the presence of integrity training alone is not necessarily enough to deter researchers from engaging in QRPs. In contrast to this finding, we find a negative association between self-reported QRP engagement and the presence of an environment that is perceived to be supportive, characterised by expertise for dealing with integrity breaches, effective mechanisms for detecting and sanctioning breaches, and protections for whistle-blowers and those accused of circumventing integrity standards. While it is not possible to disaggregate the relative role of each component, it suggests that fostering a research environment with comprehensive support for research integrity (including expertise, monitoring, and protection), beyond the mere provision of research integrity training, may be efficacious in reducing self-reported QRP engagement to some degree. It is worth nothing that this represents a researcher's perceptions of their working environment, rather than representing a true description of the institutional working environment. Additionally, individuals within the same institution might perceive conditions, policies, and structures differently. They may also have varying levels of awareness regarding the existence of these conditions, policies, and structures.

Limitations

Overall, our effect sizes are, while coherent and consistent with much of the extant literature, rather small. That said, the small effect sizes should not necessarily be interpreted as evidence that the associations are unimportant. As discussed below, self-reported QRP engagement is likely to be conservatively estimated, and selective participation may have restricted the observed variance in both self-reported QRP engagement and relevant predictors. Where outcome and predictor variability is limited, associations are likely to be correspondingly reduced. Thus, the observed effect sizes may partly reflect restriction of range rather than the absence of substantively meaningful relationships. Moreover, even small associations may be meaningful when considered at the level of the wider research population. QRPs are low-frequency but are arguably consequential behaviours, and small shifts in the probability of self-reported engagement may translate into meaningful differences when aggregated across large numbers of researchers, studies, and publications. From this perspective, modest predictor effects may still have practical significance, particularly where they identify modifiable features of the research environment or incentive structure.

Accordingly, the small effect sizes observed in this study should be interpreted cautiously and humbly but not dismissed as trivial.

A limitation of our study, not necessarily unrelated to the foregoing is that the QRPs used in the survey were selected non-randomly, and represent a distinct set of practices that qualitatively differ from some of the practices explored in other surveys. They were designed to capture behaviours that could apply equally to all kinds of research – quantitative, qualitative, and theoretical. This means that they might be regarded as a somewhat blunt tool, compared with discipline specific studies such as John et al.[1] that have only one disciplinary focus.

Furthermore, one limitation pertains to the possibility of non-response bias arising from the low response rate. This is exacerbated by the sensitive nature of some of the survey items, including QRPs. Therefore, survey engagement may be non-random, whereby researchers particularly attuned to research integrity and those more willing to engage in or disclose QRPs may be more likely to respond. Conversely, those who have engaged in QRPs, or perceived disclosure of such engagement as risky, despite assurances of anonymity, may have been less inclined to participate. As such, the sample may differ systematically from the wider population of researchers and scientists. On this basis, prevalence estimates should not be treated as precise estimates, given that they may be affected by social desirability bias and differential association. The variance of outcome variables and predictors may be attenuated by selective participation. For example, if respondents were disproportionately similar in their concern about research integrity, willingness to disclose, or exposure to research-related pressures, this may have reduced variance in self-reported QRP engagement and attenuated associations with predictor variables. On this basis, the results should be interpreted as evidence of associations within the responding sample, rather than as definitive estimates of population-level prevalence or effect sizes. Certainly, the study never set out to provide accurate estimates of prevalence, and the results should not be taken as such.

A further limitation concerns the non-causal nature of the findings. Although the models identify individual- and institution-level factors associated with self-reported QRP engagement, the cross-sectional survey design precludes conclusions about causal direction. Some associations may reflect reverse causality. For example, researchers who report engaging in QRPs may subsequently perceive their research environment as more permissive, more pressured, or less aligned with scientific norms. Alternatively, the observed associations may be non-causal, reflecting unmeasured individual, disciplinary, or institutional factors that shape both perceptions of the research environment and willingness to report QRP engagement. It is also important to emphasise that the outcome measure captures self-reported QRP engagement rather than directly observed behaviour. Responses may therefore be affected by recall error, social desirability concerns, differences in how respondents interpret QRP items, and variation in willingness to disclose sensitive behaviours. Accordingly, the results should be interpreted as associations with reported QRP engagement, rather than as evidence that the predictors cause QRP engagement.

One further limitation is that the sample focused primarily on countries in Europe, the US and Canada. On this basis, it is not possible to generalise our finding of beyond Europe and the Anglo-American sphere.

Conclusion and Future Research

Our research contributes to the explanatory framework of QRP engagement, identifying both systemic and individual-level factors that predict, albeit weakly, self-reported QRP engagement. Given the lack of institutional variance, we suspect that the system-wide pressures (to publish frequently, in high-impact journals and to obtain funding) is a major contributing factor to self-reported QRP engagement and acts directly on the individual researcher. Institutional mediation of such pressure, if it occurs, does not appear to be a strong channel. This is an interesting thought, if a little concerning, because it is easier for organisations to change than it is for systemic change to be brought about ‘top down’.

Commitment to the Mertonian norms of science seem to protect to some extent against the systemic and other individual-level influences. Studies that examine the impact of other individual-level characteristics in reducing engagement in QRPs highlight factors such as personality traits (Giluk, 2015), achievement goals (Janke, 2019), perceived ethical defensibility, and perceptions of what is considered normative behaviour in the field (Sacco et al., 2018). To further understand this and build a more comprehensive explanatory model, future research should focus on identifying other individual-level factors that might impact researcher behaviour. For example, a researcher's personal reasons behind their research decisions, such as social or political goals, financial improvement, concerns about validity and reliability, along with motivations directly linked to the scientific process like obtaining grants, increasing frequency of publication, desiring novelty, societal impact, and the replication of results, could affect their behaviour and their willingness to compromise on research integrity. A more explicit analysis of how these individual-level factors protect against, or enhance, systemic-level causes is required, and will allow for a fuller, more comprehensive explanatory model for researcher QRP engagement. This will permit the development of evidence-based interventions and policies that are efficacious in reducing QRP engagement, improving research integrity, and creating a more valid, reliable, and trustworthy science.

Availability of data and materials

The code used for the analysis in this study is openly available on the Open Science Framework: https://osf.io/sg8zf/. The primary dataset utilised for this study can be found at the

UK Data Service: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=9023. The organisational dataset analysed during the current study are not publicly available due to identifying information but are available from Nick Allum upon reasonable request.

References

6th World Conference on Research Integrity. (2019, November 23). *The Hong Kong manifesto for assessing researchers: Fostering research integrity*. World Conferences on Research Integrity. https://www.wcrif.org/downloads/main-website/hong-kong-principles/138-hkp-preprint-english/file

Allum, N., Mejlgaard, N., Sørensen, M. P., Andersen, J. P., Schneider, J. W., Charitidis, C. A., … Reid, A. (2024, February 6). International Survey on Research Integrity (IRIS) https://osf.io/xb9rk/

Allum, N., Reid, A, Bidoglia, M., Gaskell, G., Aubert-Bonn, N., Buljan, I., Fuglsang, S., Horbach, S., Kavouras, P., Marušić, A., Mejlgaard, N., Pizzolato, D., Roje, R., Tijdink, J., & Veltri G. (2023). Researchers on research integrity: a survey of European and American researchers. *F1000Research*, *12,* 187. https://doi.org/10.12688/f1000research.128733.1

Anderson, M. S. (2000). Normative orientations of university faculty and doctoral students. *Science and Engineering Ethics*, *6*, 443-61. https://doi.org/10.1007/s11948-000-0002-6

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2010). Extending the Mertonian Norms: Scientists' Subscription to Norms of Research. *The Journal of Higher Education*, *81*(3), 366-93. https://doi.org/10.1353/jhe.0.0095

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. P*erspectives on Psychological Science, 7*(6), 543-54. https://doi.org/10.1177/1745691612459060

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science: improving the standard for basic and preclinical research. *Circulation research*, *116*(1), 116-26. https://doi.org/10.1161/CIRCRESAHA.114.303819

Boumans, N. P. G., De Jong, A. H. J., & Janssen, S. M. (2011). Age-differences in work motivation and job satisfaction. The influence of age on the relationships between work characteristics and workers' outcomes. *The International Journal of Aging & Human Development*, *73*(4), 331-50. https://doi.org/10.2190/AG.73.4.d

Bouter, L. M. (2015). Commentary: Perverse incentives or rotten apples. *Accountability in Research*, *22*(3), 148-161. https://doi.org/10.1080/08989621.2014.950253

Bray, D., & von Storch, H. (2017). The Normative Orientations of Climate Scientists. *Science and engineering ethic*s, 23(5), 1351-67. https://doi.org/10.1007/s11948-014-9605-1

Brooker, R., & Allum, N. (2024). *Investigating the Links between Questionable Research Practices, Scientific Norms and Organisational Culture*. OSF. https://osf.io/sg8zf/

Chavalarias, D., Wallach, J. D., Li, A. H., & Ioannidis, J. P. A. (2016). Evolution of Reporting P Values in the Biomedical Literature, 1990-2015. *JAMA*, *315*(11), 1141-1148. https://doi.org/10.1001/jama.2016.1952

Conroy, G. (2020, February 12). *Is research integrity training a waste of time? Building good research practices begins before entering the lab*. Nature. https://www.nature.com/nature-index/news/is-research-integrity-training-a-waste-of-time

Diaz-Martinez, L. (2013, June 11). *Say 'No' to the Impact Factor*. ASCB (American Society for Cell Biology). https://www.ascb.org/careers/say-no-to-the-impact-factor/

Edwards, M. A., & Roy, S. (2017). Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental engineering science*, *34*(1), 51-61. https://doi.org/10.1089/ees.2016.0223

Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PloS one*. *4*(5), e5738. https://doi.org/10.1371/journal.pone.0005738

Fiedler, K., & Schwarz, N. (2016). Questionable Research Practices Revisited. *Social Psychological and Personality Science*, *7*(1), 45-52. https://doi.org/10.1177/1948550615612150

Francis, G. (2012) 'Publication bias and the failure of replication in experimental psychology', *Psychonomic Bulletin & Review*, 19(6), pp. 975–991. Available at: https://doi.org/10.3758/s13423-012-0322-y (Accessed: 3 June 2026).

Friesen, L., & Gangadharan, L. (2012). Individual level evidence of dishonesty and the gender effect. *Economics Letters*, *117*(3), 624-26. https://doi.org/10.1016/j.econlet.2012.08.005

Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, *145*(1), 1-44. https://doi.org/10.1037/bul0000174

Giluk, T. L., Postlethwaite, B. E. (2015). Big five personality and academic dishonesty: A meta-analytic review. *Personality and Individual Difference, 72,* 59-67. https://doi.org/10.1016/j.paid.2014.08.027

Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PLoS One*, *17*(2), e0263023. https://doi.org/10.1371/journal.pone.0263023

Higginson, A. D., & Munafò, M. R. (2016). Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. *PLoS biology*, *14*(11), e2000995. https://doi.org/10.1371/journal.pbio.2000995

Janke, S., Daumiller, M., & Rudert, S. C. (2019). Dark Pathways to Achievement in Science: Researchers' Achievement Goals Predict Engagement in Questionable Research

Practices. *Social Psychological and Personality Science*, *10*(6), 783-91. https://doi.org/10.1177/1948550618790227

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling. *Psychological Science, 23*(5), 524-32. https://doi.org/10.1177/0956797611430953

Johnsson, L., Eriksson, S., Helgesson, G., & Hansson, M. G. (2014). Making researchers moral: Why trustworthiness requires more than ethics guidelines and review. *Research Ethics, 10*(1), 29-46. https://doi.org/10.1177/1747016113504778

Jussim, L., Krosnick, J. A., Stevens, S.T., & Anglin, S. M. (2019). A social psychological model of scientific practices: Explaining research practices and outlining the potential for successful reforms. *Psychologica Belgica*, *59*(1), 353-72. https://doi.org/10.5334/pb.496

Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, *2*(3), 196-217. https://doi.org/10.1207/s15327957pspr0203_4

Kim, S. Y., & Kim, Y. (2018). The Ethos of Science and Its Correlates: An Empirical Analysis of Scientists' Endorsement of Mertonian Norms. *Science, Technology and Society*, *23*(1), 1-24. https://doi.org/10.1177/0971721817744438

Kornfeld, D. S. (2012). Perspective: research misconduct: the search for a remedy. *Academic medicine: journal of the Association of American Medical Colleges*, *7*(7), 877-82. https://doi.org/10.1097/ACM.0b013e318257ee6a

Kretser, A., Murphy, D., Bertuzzi, S., Abraham, T., Allison, D. B., Boor, K. J., Dwyer, J., Grantham, A., Harris, L. J., Hollander R., Jacobs-Young, C., Rovito, S., Vafiadis, D., Woteki, C., Wyndham, J., & Yada, R. (2019). Scientific Integrity Principles and Best

Practices: Recommendations from a Scientific Integrity Consortium. *Science and engineering ethics*, *25*(2), 327–355. https://doi.org/10.1007/s11948-019-00094-3

Latan, H., Chiappetta Jabbour, C. J., Lopes de Sousa Jabbour, A. B., & Ali, M. (2021). Crossing the red line? Empirical evidence and useful recommendations on questionable research practices among business scholars. *Journal of Business Ethics, 184*, 549–569. https://doi.org/10.1007/s10551-021-04961-7

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), 424-32. https://doi.org/10.1177/1745691613491437

MacCoun, R. J. (1998). Biases in the interpretation and the use of research results. *Annual Review of Psychology*, *49*, 259-87. https://doi.org/10.1146/annurev.psych.49.1.259

MacCoun, R. J. (2022). P-hacking: A strategic analysis. In Jussim L. J., Krosnick, J. A., Stevens, S. T., (Ed.). *Research integrity: Best practices for the social and behavioral sciences* (pp. 295-313).Oxford University Press. https://doi.org/10.1093/oso/9780190938550.003.0011

Macfarlane, B., & Cheng, M. (2008). Communism, Universalism and Disinterestedness: Re-examining Contemporary Support among Academics for Merton's Scientific Norms. *Journal of Academic Ethics*, *6*(1), 67-78. https://doi.org/10.1007/s10805-008-9055-y

Maggio, L., Dong, T., Driessen, E., & Artino A. Jr. (2019). Factors associated with scientific misconduct and questionable research practices in health professions education. *Perspectives on medical education*, *8*(2), 74-82. https://doi.org/10.1007/s40037-019-0501-x

Martinson, B. C., Anderson, M. S., & de Vries, R. (2005). Scientists behaving badly. *Nature*, *435*(7043):737-38. https://doi.org/10.1038/435737a

Martinson, B. C., Crain, A. L., Anderson, M. S., & De Vries, R. (2009). Institutions' expectations for researchers' self-funding, federal grant holding, and private industry involvement: manifold drivers of self-interest and researcher behavior. *Academic medicine: journal of the Association of American Medical Colleges*, *84*(11), 491-99. https://doi.org/10.1097/ACM.0b013e3181bb2ca6

Merton, R. K. (1979). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.

Moher, D. (2018). Reporting guidelines: doing better for readers. *BMC medicine, 16*(1), 233. https://doi.org/10.1186/s12916-018-1226-0

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., Simonsohn U, Wagenmakers, E. J., Ware, J. J., Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature human behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Necker, S. (2014). Scientific misbehavior in economics. *Research Policy*, *43*(10), 1747-59. https://doi.org/10.1016/j.respol.2014.05.002

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, *7*(6), 615-631. https://doi.org/10.1177/1745691612459058

Nuzzo, R. (2015). How scientists fool themselves - and how they can stop. *Nature*, *526*(7572), 182-5. https://doi.org/10.1038/526182a

Phillips, T., Nestor, F., Beach, G., & Heitman, E. (2018). America COMPETES at 5 years: an analysis of research-intensive universities' RCR training plans. *Science and Engineering Ethics*, *24*(1), 227-49. https://doi.org/10.1007/s11948-017-9883-5

Pupovac ,V., & Fanelli D. (2015). Scientists Admitting to Plagiarism: A Meta-analysis of Surveys. *Science and engineering ethics*, *21*(5), 1331-52. https://doi.org/10.1007/s11948-014-9600-6

Resnik, D. B. (2011). Scientific research and the public trust. *Science and engineering ethics*, *17*(3):399-409. https://doi.org/10.1007/s11948-010-9210-x

Robinson, R. J., Lewicki, R. J., & Donahue, E. M. (2000). Extending and testing a five factor model of ethical and unethical bargaining tactics: Introducing the SINS scale. *Journal of Organizational Behavior*, *21*(6), 649-64. https://www.jstor.org/stable/3100394

Rubin, M. (2023) 'Questionable metascience practices', *Journal of Trial & Error*, 4(1), pp. 5–20. Available at: https://doi.org/10.36850/mr4 (Accessed: 3 June 2026).

Rubin, M. (2025) *A brief review of research that questions the impact of questionable research practices*. PsyArXiv. Available at: https://doi.org/10.31234/osf.io/ah9wb_v3 (Accessed: 3 June 2026).

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In Defense of the Questionable: Defining the Basis of Research Scientists' Engagement in Questionable Research Practices. *Journal of empirical research on human research ethics: JERHRE*, 13(1), 101-10. https://doi.org/10.1177/1556264617743834

Samota, E. K., & Davey, R. P. (2021). Knowledge and Attitudes Among Life Scientists Toward Reproducibility Within Journal Articles: A Research Survey. *Frontiers in research metrics and analytics*, *6*, 678554. https://doi.org/10.3389/frma.2021.678554

Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(36), 15268-15273. https://doi.org/10.1073/pnas.0907352106

Satalkar, P., & Shaw, D. (2019). How do researchers acquire and develop notions of research integrity? A qualitative study among biomedical researchers in Switzerland. *BMC Medical Ethics*, *20, Article* 72. https://doi.org/10.1186/s12910-019-0410-x

Schneider, J. W., Allum, N., Andersen, J. P., Petersen, M. B., Madsen, E. B., Mejlgaard, N., & Zachariae, R. (2024). Is something rotten in the state of Denmark? Cross-national evidence for widespread involvement but not systematic use of questionable research practices across all fields of research. *PloS one*, *19*(8), e0304342. https://doi.org/10.1371/journal.pone.0304342

Simmon, 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), 1359-66. https://doi.org/10.1177/0956797611417632

Tijdink, J. K., Bouter, L. M., Veldkamp, C. L., van de Ven, P. M., Wicherts, J. M., & Smulders, Y. M. (2016). Personality Traits Are Associated with Research Misbehavior in Dutch Scientists: A Cross-Sectional Study. *PloS one*, *11*(9), e0163251. https://doi.org/10.1371/journal.pone.0163251

Tijdink, J. K., Vergouwen, A. C. M., & Smulders, Y. M. (2013). Publication pressure and burn out among dutch medical professors: A nationwide survey. *PLOS One*, *8*(9), e73381. https://doi.org/10.1371/journal.pone.0073381

Ulrich, R. and Miller, J. (2020) 'Questionable research practices may have little effect on replicability', *eLife*, 9, e58237. Available at: https://doi.org/10.7554/eLife.58237 (Accessed: 3 June 2026).

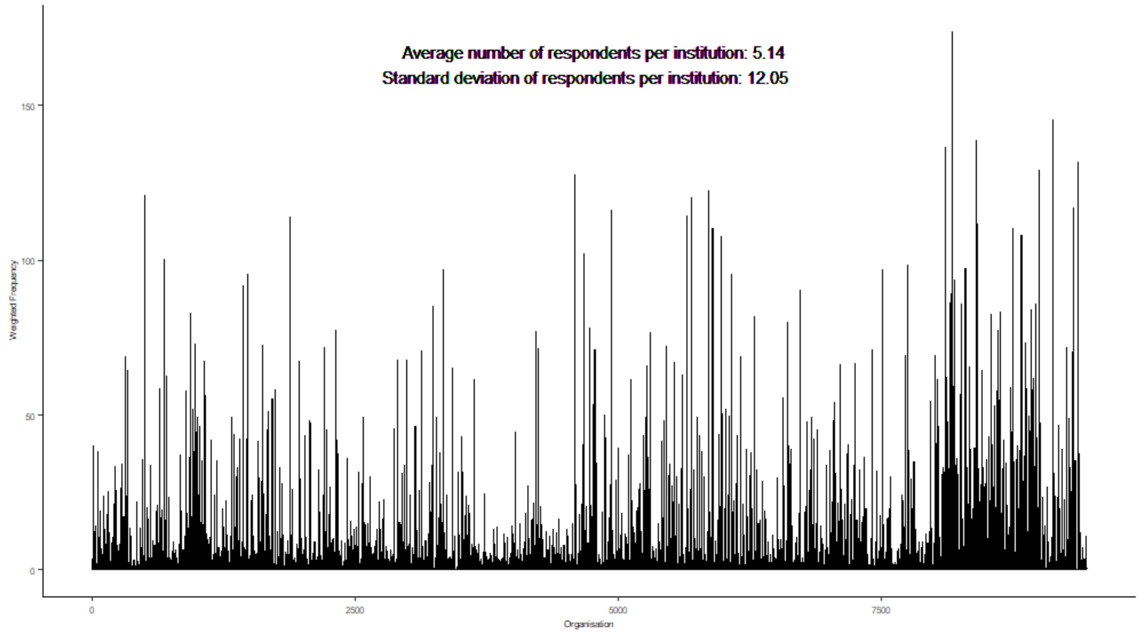
Ware, J. J., & Munafò, M. R. (2015). Significance chasing in research practice: causes, consequences and possible solutions. *Addiction*, *110*(1), 4-8. https://doi.org/10.1111/add.12673

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, 7, 1832. https://doi.org/10.3389/fpsyg.2016.01832

Xie, Y., Wang, K., & Kong, Y. (2021). Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis*. Science and engineering ethics, 27*(4), 41. https://doi.org/10.1007/s11948-021-00314-9

Supplementary Material

Table S2.1. Weighted frequency of respondents by organisation (n = 7,666)



Chapter 3

The influence of subjective beliefs about others' behaviour on research practice: results from an empirical analysis

Abstract

A growing literature has sought to identify how to best tackle the ‘replication crisis’. Some commentators advocate for better training in research design, while others propose an overhaul of incentive structures. We explore a potential further means of counteracting practices that produce non-replicable findings by investigating a mechanism that potentially influences the decision to engage in QRPs: researchers’ subjective beliefs about QRP prevalence. These beliefs may create ‘descriptive norms’ (i.e., what is perceived as typical) pointing towards a prevalence that undermine ‘injunctive norms’ (what is perceived as ethically appropriate) nurturing good research practice (Fiedler and Schwarz, 2016). A belief that QRPs are prevalent may contribute to a perceived necessity to engage in such behaviours in order to keep up with peers. We empirically investigate this mechanism using data collected as part of the ‘Practices, Perceptions, and Patterns of Research Integrity’ (PRINT) project, from a survey administered in 2018. The sample (N=3402) consisted of post-doctoral researchers working at 18 universities across Europe and the US and represent all scientific fields. Respondents reported estimates of prevalence in recent literature, their own behaviour and perceived damage to science for a range of QRPs. First, we report the associations between perceived prevalence and self-reported engagement in QRPs and explore how these vary across fields and career-stage. Then, we report findings from a series of mediation and moderation analyses that suggest plausible mechanisms and conditions under which this relationship occurs. We conclude by discussing key implications for improving research quality.

Key Words: Questionable research practices, research integrity descriptive norms, injunctive norm

Introduction

Over the past two decades there has been growing concern over the credibility and reproducibility of findings in the social and natural sciences. This concern has been fuelled by empirical research demonstrating that a significant quantities of published research findings are false (Ioannidis, 2005; Johnson, 2013). Doubts over the veracity of published research findings have been elevated by numerous highly publicised replication attempts that have failed to reproduce significant quantities of published findings, most notably in psychology and medicine (Open Science Collaboration, 2015; Camerer et al., 2016, 2018). For example, the Open Science Collaboration (2015) found that only 39% of findings from 100 experimental and correlational studies published in three high-ranking psychology journals were replicable (cf. Gilbert et al., 2015). Similarly, Camerer et al. (2016) identified statistically significant effects in the *same direction* as those found in the original study in only 11 of 18 experimental economic studies replicated from two high-ranking journals.

This 'reproducibility crisis' has less to do with overt research misconduct, typically defined in terms of falsification, fabrication and plagiarism (FFP), and more to do with 'questionable research practices' (QRPs; John et al., 2012; Simmons et al., 2011).[7] QRPs are without straightforward definition, but the term broadly refers to practices that exist in an ambiguous 'grey area' between sound science and outright scientific misconduct. Examples include developing hypotheses after results are known (termed "HARKing"; Kerr, 1998, including authors on publications who have not made significant contributions to the work to warrant inclusion, and failing to report the results of all measures from an experiment. QRPs

---

[7] However, it is necessary to note that the extent of the impact of QRPs on replication is contested (e.g., Ulrich & Miller, 2020; Rubin, 2023; Rubin, 2025; Francis, 2012). In fact, the extent to which replication rates are in fact a problem has been subject to refute (Rubin, 2023; Rubin, 2025; Fanelli, 2018; Oberauer & Lewandowsky, 2019). That said, it is important to note that, while sometimes QRPs are justifiable in particular contexts, they still have important downstream consequences, including the production of false research literatures, wasted resources (including time and money) and reduced credibility of legitimate findings. Therefore, QRPs are not treated merely as contributors to the replication crisis, but as practices that undermine the integrity and reliability of research more broadly.

involve the exploitation of the inherent flexibility that exists in the research process. There is the potential for such practices to be utilised for the purpose of presenting biased evidence in support of conclusions (Banks et al., 2016). However, given their inherent ethical ambiguity, these practices may be seen as warranted or defensible in certain contexts (John et al., 2012; Sacco et al., 2018) and may be engaged in without intent to deceive and based on normative assumptions that prevail within a particular research field or subfield.

Methodological decisions (so-called, "researcher degrees of freedom") that exist at various stages of the research process can increase Type I error rates (Simmons et al., 2011 1359). For example, including theoretically irrelevant covariates in statistical models, analysing two dependent variables without adjusting alpha values or excluding data points based on post-hoc criteria can increase the probability of obtaining false-positive findings. Relatedly, failure to publish non-significant findings (i.e., "the file drawer problem") as a result of bias in the selection process or pressure from sponsors, increases the likelihood that published results will reflect Type-I errors, biasing effect sizes upwards in meta-analyses (Franco et al., 2014). Other QRPs can lead to different problems. For example, including authors on a paper where the contributions made do not warrant authorship inflate the perceived expertise and competence of that individual, bestowing them with an unfair advantage in hyper-competitive academic environments (e.g., when seeking tenure or promotion).

QRPs are of particular concern given the frequency at which researchers engage in them. While only 1.97% of researchers admit to having engaged in research misconduct (i.e., fabrication, falsification, and data modification), meta-analytical research suggests that 33.7% admit to having engaged in QRPs (Fanelli, 2009 cf. Xie et al., 2021; Fiedler & Schwarz, 2016). When asked about the behaviour of their colleagues, admission rates were significantly higher at 72%. However, given the unethical nature of the behaviour, respondents are less likely to provide honest reports of own behaviour, therefore estimates of prevalence likely underrepresent true rates of engagement. Addressing this, John et al. (2012) using a scoring-algorithm to incentivise truth-telling, found that 94% of respondents admitted to having engaged in at least one QRP. This high frequency suggests that the validity of large quantities of published research findings may have been compromised.

On this basis, numerous interventional and educational strategies have been proposed as a means of reducing the prevalence of QRPs, and hence improving the quality of scientific research. One strategy focuses on enhancing the quality and frequency of training in research ethics (Johnson et al., 2014 Holm & Hofmann, 2018) and research methodology (Munafò et al., 2017). This strategy is predicated on the belief that QRPs, at least in part, result from researchers' lack of awareness of how to appropriately conduct research and avoid suboptimal research practices. For example, while a majority of researchers admit to having engaged in behaviours that are best described as QRPs, most believe in the integrity of their own research and the appropriateness of their own practices (John et al., 2012).

Other observers emphasise the misalignment between incentive structures and the ideals of science, and this is the pre-eminent theory explaining why researchers compromise their integrity. Academic institutions often use simple metrics, such as frequency of publication and journal impact factors, when making assessments for hiring, tenure and promotion (Moher et al., 2018). This is accompanied by a publication bias that favours the aesthetic quality of results (e.g., novelty and statistical significance; Sterling, 1959; Dwan et al., 2013; Chavalarias et al., 2016). For example, one study found that statistically significant results (i.e., p-value <0.05) account for approximately 96% of published abstracts in MEDLINE and full-text articles in PubMed Central (Chavalarias et al., 2016). Therefore, institutional pressures to publish, combined with a publication bias, potentially incentivise researchers to engage in suboptimal research behaviours, such as selective reporting, in order to increase the likelihood of yielding statistically significant results (Head et al., 2015; Simmons et al., 2011).

Additionally, direct relationships between publication pressure and suboptimal research behaviour have been found (Tijdink et al., 2014; Paruzel-Czachura et al., 2020). However, beyond publication pressure and statistical and methodological ignorance, there are likely other factors that motivate, and in some cases necessitate, engagement in QRPs. In the current study, we seek to identify one of these facilitating factors: researcher's 'perceptions' of QRP prevalence.

Perceptions of Norms, Competition and Research Behaviour

Social psychological literature distinguishes between norms that describe typical behaviours (so-called, 'descriptive norms') and desirable behaviours (define as 'social norms' by Bicchieri (2017) and 'injunctive norms' by Cialdini et al. (1991)). Descriptive norms (i.e., 'what most people do' (Cialdini et al., 1991)) have been found to influence behaviours in various contexts in the direction of the norm (Cialdini et al., 2006; Goldstein et al., 2008; Gerber & Rogers, 2009; Stangor et al., 2001; Schultz et al., 2007) and have been used in many behavioural change interventions, often in combination with injunctive norms (Bertholet et al., 2016; Perkins et al., 2011). As well as motivating pro-social behaviour, descriptive norms have been found to motivate engagement in unethical practices, including academic cheating (Smyth & Davis, 2004; Chapman & Lupton, 2004; Stone et al., 2009) and research misconduct (Sacco et al., 2017; Rajah-Kanagasabai & Roberts, 2015). The objective accuracy of these perceptions is less important than their subjective influence on researchers' actual behaviours (Tankard & Paluck, 2016). Nevertheless, estimates of QRP prevalence do seem to reflect genuinely high engagement rates in these practices (Fanelli, 2009). This aligns closely with Thomas and Thomas's (1928, pp. 571–572) classic theorem: "If men define situations as real, they are real in their consequences."

Perceiving QRPs as descriptively normative may motivate researchers to engage in those behaviours through processes of norm conformity (Asch, 1956; Sherif, 1936) and informational dependence. Indeed, where behaviours are ethically ambiguous, individuals are more likely to look to others for guidance on how to behave (Festinger, 1954). Further, perceiving QRPs as prevalent may feasibly influence a researcher's decision to engage in those practices, through increasing competitive norms that compound pressures to cut corners and cheat in order to keep up with peers. Certainly, the feeling of being disadvantaged by abiding by research standards that others circumvent was seen as a cause of engagement in QRPs by a majority of social psychologists in Stürmer et al.'s (2017) study. This finding aligns with research demonstrating that in contexts marked by social ambiguity, beliefs about others' cooperative intent predict cooperative behaviour (Anderson et al., 2004), while competition incentivises individuals to improve their position relative to their competitors, often through use of illegitimate and unethical means (Belot & Schröder, 2013).

Competition is ubiquitous in the scientific system. Performance metrics in institutional assessments for hiring, tenure and promotion contribute to increased value placed on improving performance ('a unidirectional drive upwards' in Festinger's (1954) social comparison theory), which when placed in the social context facilitate interpersonal social comparison and fierce competition. The scarcity of academic jobs relative to the increasing number of PhDs (Larson et al., 2014) encourages individuals to compete for a limited number of positions, while scientists are required to compete with others for limited extramural funding. Moreover, there is a drive in some scientific fields to be the first to make a discovery, motivating researchers and research teams to compete for the recognition and prestige by 'finding it first' (Merton, 1957). Individuals are less concerned about ethical standards in competitive situations (Butterfield et al., 2000). When competition heightens perceived rivalry between individuals it decreases the moral costs of transgressing norms by using unethical means to beat competitors (To et al., 2020). Certainly, descriptive norms suggesting that unethical behaviours are common have been found to undermine injunctive norms fostering ethical behaviour in certain contexts (Cialdini et al., 1990; Staunton et al., 2014). On this basis, research environments are likely fertile grounds for the influence of descriptive norms on behaviour, whereby perceiving others to be engaging in QRPs necessitates engagement in those behaviours for sake of obtaining a competitive advantage or safeguarding one's career.

The Current Study

Preliminary evidence for the influence of descriptive norms on research behaviour is derived from Sacco et al. (2018) study of principal investigators funded by the National Institutes of Health in the United States. It was found that the more respondents believed QRPs to be normative, the more ethically defensible QRPs were perceived to be, and the more willing researchers were to engage in them. In the current research, we build on these findings by seeking to identify whether perceptions of prevalence are positively associated with the extent to which researchers have engaged in QRPs, across all major fields of research, using an international sample. As such, utilising data from a cross-sectional survey, we seek to test the following primary hypothesis:

> H1: The more prevalent a researcher perceives QRPs to be within their scientific field, the greater the extent of their self-reported own use of those practices.

Judging suboptimal research conduct to be common within a scientist's disciplinary field potentially contributes to competitive norms that motivate researchers to cut corners in order to keep up with colleagues, thereby preserving their career. As such, researchers may engage in QRPs out of a perceived career necessity. Certainly, following To et al. (2020), perceiving a behaviour to be prevalent, maximises norms of competitiveness, thus decreasing the moral costs associated with engaging in unethical behaviour. On this basis, we undertake a mediation analysis seeking to test whether perceptions of prevalence have an indirect effect on the extent of self-reported engagement in QRPs by increasing the strength of belief that compromising research integrity is necessary for the sake of one's career. On this basis, we test the following hypothesis:

> H2: Estimates of the extent of QRP prevalence within a researcher's field have an indirect effect on the degree to which the researcher engages in those practices through increasing the degree to which it is believed that not compromising research integrity will have a negative impact on one's career.

> Perceptions of the behaviours of other scientists are likely to have greater influence on a scientist's own conduct when the behaviour is seen as damaging to science. However, if a QRP is judged to be undamaging to science, scientists are likely to engage in the behaviour regardless of the perceived conduct of others. Thus, we seek to test the following hypothesis:

> H3: The relationship between perceptions of prevalence of QRPs within a researcher's field and the extent of their own use of those practices will be greater when the practices are seen as more damaging to science.

The relationship between perceptions of prevalence and extent of self-reported engagement in QRPs is feasibly affected by the researcher's stage of career. On the one hand, the longer a researcher has worked in the field, the more likely they are to have adjusted their behaviour and adopted practices that align with the normative framework of that group. On the other hand, early career researchers (ECRs) may be more vigilant of the behaviour of others and have more to lose for not conforming to established normative frameworks that pervade the discipline and are therefore more likely to be influenced by them. On this basis, we test the following non-directional hypothesis:

H4: The relationship between perceptions of prevalence of QRPs within a researcher's scientific field and the degree to which the researcher has used those practices varies according to the researchers' career stage.

There are feasibly scientific field differences in the relationship between prevalence estimates and extent of self-reported engagement in QRPs. However, we have no a priori beliefs as to whether a researcher's scientific field moderates this relationship, thus we include it as a planned exploratory analysis, stratifying this relationship by the researchers reported disciplinary field.

Finally, following Fanelli (2009) researchers are more likely to perceive QRP prevalence to be greater when assessing the behaviour of others, as opposed to when making judgements of their own behaviour. Therefore, we test the following hypothesis:

H5: Researchers perceive the prevalence of QRPs to be greater in their scientific field compared to in their own sole or co-authored publications.

Methodology

Pre-registration

Research questions and hypotheses were pre-registered with the Center for Open Science in May 2021 (https://shorturl.at/QuqM2).

Data

The study utilises secondary data from a large-scale cross-sectional international survey administered as part of the Practices, Perceptions, and Patterns of Research Integrity project in 2018 (PRINT). The PRINT project was anchored at the Danish Centre for Studies in Research Policy (CFA), Aarhus University. The survey protocol was pre-registered with the Centre for Open Science in October 2018 (available, along with reproducibility materials at Schneider, 2018).

Sample

The survey was administered to researchers above PhD level at eighteen universities (approximately 55,000 researchers) in Denmark, Croatia, the United Kingdom, Austria and the United States. Respondents represented all major fields. Webscraping was used to obtain contact information for the target population and respondents were invited to participate through email. The final population sample consisted of 4712 respondents. Due to violations of the Missing Completely at Random (MCAR) assumption, missing values were addressed using the mice package in R (version 4.0.3; Van Buuren & Groothuis-Oudshoorn, 2011), resulting in 50 imputed datasets. QRP items that were missing by-design were not imputed, and all survey items were included in the imputation model. Respondents who identified as non-binary on the gender item (n = 55) were excluded listwise due to small cell sizes. Additionally, miscoded responses for scientific discipline (n = 11) and country (n = 2) were also removed listwise. Consequently, the final analytical sample comprised 4,646 respondents.

Table 3.1 presents the characteristics of these respondents, stratified by primary disciplinary field, country, and gender. Descriptive statistics are based on the first-imputed dataset.

Table 3.1. Characteristics of sample by primary disciplinary field, country and gender

	Disciplinary Field				Total
	Social Sciences	Natural and Technology Sciences	Arts and Humanities	Medical and Health Sciences	Total Sample
Female (%)	8.1	10.8	6	11	36.9
Male (%)	11.9	30.3	7	14.9	64.1
Denmark (%)	62.2	74.4	70.1	77.9	72.3
Croatia (%)	8.5	7.3	5.6	4.7	6.7
UK (%)	8.9	7.8	4.7	5.8	7.1
US (%)	4.5	2.1	3	3	2.9
Austria (%)	15.9	8.4	16.6	8.6	11
Overall (%)	20	41.2	13	25.9	

Variables

Questionable Research Practices (QRPs). A total of 25 research behaviours pertaining to data analysis strategy and scientific reporting that are widely considered to be 'questionable' were generated for the purposes of the survey. These behaviours are listed in Table 3.2. Each respondent was provided with 9-10 of these QRPs, allocated randomly, with each item accompanied by three discrete questions assessing a) the extent to which the respondent believes the practice is used in recent publications within their field of research (1 = *In no recent publications*, 7 = *In all recent publications*), b) the extent to which the respondent has used these practices in recent sole or co-authored publications (1 = *In no recent publications*, 7 = *In all recent publications*) and c) the extent to which the respondent believes that this practice is damaging the trust in research findings and claims in their field of research (1 = *To a very small extent*, 7 = *To a very great extent*). No intermediate labels were used.

Higher scores indicated a greater perception of prevalence, a greater extent of past self-reported QRP engagement and a stronger belief that the practice is damaging to trust in research findings and scientific claims. By anchoring the self-reported engagement to recent timeframes and specific output contexts rather than lifetime exposure, the measure provides a more localised behavioural snapshot and implicitly account for differing disciplinary productivity rates. This avoids inflated and misleading estimates that would occur with measures that are unbounded and measure lifetime exposure (an issue highlighted by Fiedler & Schwarz, 2016).

Table 3.2. Questionable research practice items included in the PRINT survey

QRP Number	Questionable Research Practice (QRP) Item Wording
1	Including authors on a publication who has not contributed sufficiently to the work to merit authorship.
2	Failing to offer authorship to collaborators who had contributed sufficiently to the work to merit authorship
3	In a publication, failing to disclose relevant personal, financial, political or intellectual conflicts of interests.
4	In significance testing studies, continue to collect more data in order to render non-significant results significant. For example: If initial results are not statistically significant, collect more data until you get the desired result.
5	In significance testing studies, continue to reanalyse data until a statistically significant result is obtained. For example: If initial results are not statistically significant, remove or transform data; use other statistical tests; redefine the outcome variable, switching to an alternate control group; trying various combinations of independent and control variables; analysing various subgroups and so forth; thus, reanalysing data until you get the desired result.
6	Publishing work despite knowing that it is redundant and does not contribute to the existing knowledge base. Please note, reviews, translations and other types of research publications are not necessarily redundant if they fill a scholarly need in the scientific communication
7	Citing literature deemed relevant for your study without actually having read it.
8	Claim to have used a particular qualitative analytical approach appropriately, for example “grounded theory” or “triangulation”, when this was not the case.
9	Avoiding to share data, research protocols, information on experimental setup, instrumentation, coding, or other information about a study requested by colleagues to evade transparency.
10	Agree to review a manuscript despite knowing that you have inadequate expertise to provide a competent review.

11	When reviewing a manuscript, not investing the effort necessary to conduct a thorough review.
12	Submitting a biased review report that evaluated the manuscript unfairly.
13	In a publication, deliberately reuse all or parts of previously published data without disclosure.
14	Deliberately divide the results of a study over more publications than needed with the intention to increase the number of publications.
15	Selectively citing irrelevant or unnecessary publications to please reviewers or editors.
16	Deliberately cite own publications more than warranted by their relevance, to promote the visibility of your work or improve your citation metrics.
17	Deliberately disregard citing relevant publications that contradict own beliefs or research.
18	Selectively focusing on parts of the data or source material that support your preconceptions or hypotheses and deliberately disregarding parts that do not. Source material can be of all kinds, such as interviewees, documents, primary sources, etc
19	Deliberately refrain from reporting findings that could weaken or contradict own theories, hypotheses or findings
20	Wilfully present findings as more "clear-cut" than justified by the data. For example, neglecting to disclose contradictory results, or thoroughly discuss study limitations, or deliberately overlooking counter arguments
21	Without disclosure, formulate or change hypotheses after having seen the results, thereby presenting an unexpected finding as having been predicted from the start in the form of a research hypothesis (quantitative study).
22	Without disclosure, presenting an unexpected finding as having been predicted from the start (qualitative study).
23	Present statistically significant main findings without distinguishing between their "statistical significance" and their potential practical or theoretical importance. For example, claiming a "significant" finding solely based on the p-value and not the effect size. Notice, importance refers to clinical, economic, biological, psychological, sociological etc
24	Report main findings that turned out not to be "statistically significant", as evidence for no difference, no effect or no association between study groups or variables.
25	Deliberately using another researcher's unpublished idea without giving credit. For example, publishing an idea voiced by a colleague at an informal meeting without giving her/him credit.

Research examining the relationship between explanatory factors and self-reported QRP engagement typically computes these variables as average scores derived from respondents' mean responses across all QRP items (e.g., Brooker & Allum, 2024;

Gopalakrishna et al., 2023; Sacco et al., 2017). Such an approach reflects the 'average' effect of a factor – such as average perceived prevalence – on average self-reported QRP engagement. However, our primary hypothesis (H1) specifically investigates the direct, measurement-level relationship between perceptions of the prevalence of individual QRPs (QRPs 1...k) and self-reported engagement in those same QRPs. That is, we aim to determine whether perceiving a particular QRP as more prevalent is directly associated with increased personal self-reported engagement in *that* QRP. Analysing average scores alone can obscure meaningful measurement-level variability. For instance, two respondents' identical average scores for QRP self-reported engagement could result from consistently moderate self-reported engagement across all QRPs or from highly variable engagement, with some QRPs rarely practiced and others frequently practiced. Consequently, using average scores alone results in information loss about the specific relationship between perceived prevalence and individual QRPs.

To address this, responses to QRP items are conceptualised as repeated measurements nested within respondents. The data were restructured into a hierarchical (long-form) format with two levels: individual measurement occasions (Level 1; n = 41,756) clustered within respondents (Level 2; n = 4,646). Items measuring perceived prevalence and self-reported engagement were cluster-mean centered and standardised within individuals, thereby isolating the within-person variance. This procedure enables the estimation of a precise, measurement-level relationship between perceptions of QRP prevalence and actual self-reported QRP engagement.

Additionally, we computed overall summary variables capturing respondents' general perceptions of how frequently other researchers engage in QRPs and their overall degree of QRP self-reported engagement. For each respondent, mean scores for these variables were derived from their responses across the 9 – 10 individual QRP items. These overall scores were then recoded into two seven-point ordinal categorical variables using equal-width binning, aligning them with the original item response scale. Higher scores indicate greater perceived prevalence of QRPs among researchers and higher overall personal self-reported engagement in QRPs.

Respondent-level Measures. Each respondent provided their main field of scientific study and were grouped into four categories: social sciences (n=931), natural and technical sciences (n=1912), medical and health sciences (n=1201) and arts and humanities (n=602). This variable was recoded into three (K-1) indicator variables representing each scientific field, with 'social sciences' as the reference category. The social sciences were used as the reference category as they occupy an intermediate position between the natural/technical sciences, the medical/health sciences, and the arts and humanities. This makes them a useful baseline for highlighting disciplinary contrasts in self-reported QRP engagement. Respondents additionally reported the number of years that have passed since completing their PhD. This item was recoded into a single indicator variable, with ECR (0-8yrs since completing their PhD) as the reference category, and mid-late career researcher (9+ years) coded as 1. This categorisation reflects common practice in the literature on research careers and integrity, where analyses often contrast early-career and more established researchers, rather than treating years since PhD as a continuous measure. A single item was used to measure the extent to which respondents viewed it as necessary to compromise research integrity for the sake of one's career. Respondents were asked the extent to which they agree with the question 'If I conduct my research without compromising research integrity, it will have negative effects on my career' (0 = *strongly agree*, 7 = *strongly disagree*).

Controls. We include several respondent-level covariates to increase precision in estimates of primary parameters of interest. Each respondent reported their gender. This variable was recoded into an indicator variable representing the gender of the respondent, with 'female' as the reference category. Males have been found to be overrepresented in cases of research misconduct (Fang et al., 2013 cf. Fanelli et al., 2015). On this basis, it is included as a control variable in models containing other respondent-level covariates. We additionally computed four (K-1) indicator variables representing the respondent's country of employment, with the 'United Kingdom' as the reference category. Finally, we test to see whether self-reported engagement in research integrity training is associated with less self-reported engagement in QRPs. Past research suggests that research integrity training at best has weak associations with researcher conduct (see, Anderson et al., 2007; Antes et al., 2010). On this basis, we create a single binary indicator variable, with 0 = no research integrity training and 1 = have engaged in research training.

Analysis Strategy

To account for the non-independence of responses to the QRP items within each respondent, and in recognising the hierarchical nature of the data, we use a Bayesian multilevel modelling framework, implemented using the Stan programming language. Specifically, we use the brms interface (Version 2.16.1; Bürkner, 2017) to Stan in R. Stan is a probabilistic programming language that implements full Bayesian statistical inference using Markov Chain Monte Carlo (MCMC) sampling, using the No-U-Turn Sampler (NUTS). Bayesian modelling allows for the derivation of a joint posterior distribution for model parameters based on a prior probability distribution and likelihood function. The posterior distribution expresses the uncertainty in the parameters given the observed data. It combines prior beliefs about parameters with the data via the likelihood, yielding updated uncertainty in the parameter values. That is, it affords a more natural quantification of the uncertainty in inferences (Kruschke, 2014), thus quantifying support in favour of null and alternate hypotheses. Multilevel modelling allows us to incorporate all information contained in the data into the model, while controlling for dependencies between multiple responses from the same respondent. Using a hierarchical specification induces partial pooling, shrinking measurement-level estimates toward the group-level distribution and balancing pooled and unpooled information.

The regression modelling strategy is performed as follows. In Model 1 (Equation 1) we specify a Bayesian logistic regression model, regressing overall QRP on overall estimates of QRP prevalence. This allows us to derive the posterior distribution of all credible parameter values indicating the strength of the relationship between respondent's overall estimate of prevalence and overall self-reported engagement in QRPs. Appropriate priors are placed on α and γ. Formally:

$$\Pr(Z_i = 1) = \text{logistic}(\alpha + \gamma \cdot \text{PrevalenceEstimate}_i).$$

(Equation 1)

In Model 2 (Equation 2) we specify a completely pooled Bayesian cumulative (proportional odds) ordered logistic regression model, whereby our primary dependent variable, extent of self-reported QRP engagement, is regressed on the primary measurement level predictor, estimates of QRP prevalence, as well as perceived damage to trust in scientific findings. $Y_i \in \{1, ..., K\}$ denotes the observed ordinal self-reported engagement level for respondent *i*, arising from discretising an underlying latent continuous variable Y_i^*. The cumulative probabilities are modelled as:

$$\Pr(Y_i \le k) = \text{logistic}(\theta_k - \eta_i)$$
$$\eta_i = \beta_1 \cdot \text{PrevalenceEstimate}_i + \beta_2 \cdot \text{TrustDamage}_i$$

(Equation 2)

Ordered thresholds $\theta_1 < \cdots < \theta_{K-1}$ and priors on β_1 and β_2 (and on the θ_k that enforce ordering). This specification reflects the assumption that the observed ordinal categories are discretised observations of a latent propensity to engage in QRPs; the logistic CDF provides the link function. Ordinal models are deemed appropriate given that the dependent variable cannot be assumed to be continuous (see, Kruschke & Liddell, 2018). Moreover, use of ordered categorical models affords us the technical advantage of avoiding strong scaling and normality assumptions. We evaluate the proportional odds assumption and conduct posterior predictive checks to assess model fit.

In Model 3 we specify a Bayesian multilevel ordinal logistic model with measurement-level independent variables and the primary independent variable. In Model 4 we introduce respondent-level independent variables, demographic controls (to increase precision), and a measurement-level interaction term for estimates of the extent to which the QRP damages scientific findings and prevalence estimates. We additionally introduce cross-level interactions between prevalence estimates and (a) scientific discipline and (b) career-stage. Parameters for measurement-level independent variables that do not vary substantially between respondents in Model 3 will be included in Model 4 as fixed effects only. Estimating category-specific effects with a cumulative model can lead to invalid category probabilities, thus causing model fitting issues (Bürkner & Vuorre, 2019). To see if improved fit is due to the model class rather than those flexible effects, we also fit an adjacent-category model without category-specific effects.

We specify minimally informative priors for all population-level parameters, assigning independent Cauchy prior probability distributions on the log-odds scale, centered at zero with scale 0.5 for metric predictors and 0.4 for binary predictors. The heavy tails that accompany the *t* family of distributions allow for robust inference (Berger & Berliner, 1986; Lange et al., 1989). These priors put little weight on implausibly large effects, while still allowing for more moderate, data-driven effects. That is, the use of the Cauchy distribution and minimally informative priors regulate extreme inferences, thus facilitating efficient computation (Gelman et al., 2008). A series of prior predictive checks and sensitivity analyses were undertaken. Alternative reasonable prior choices did not materially alter substantive inferences. The regression models were fit to each of the 50 imputed datasets, and the resulting posteriors were combined into a single inference by concatenating posterior draw across imputations to form a single combined posterior distribution. Models were additionally fit to a complete-case dataset (listwise deletion) to determine robustness and sensitivity of estimates to missing data.

We specify two Bayesian ordinal logistic mediation models (Models 5A and 5B) with cumulative link functions to test H2. That is, the indirect association between overall perception of QRP prevalence and overall self-reported engagement in QRPs through the extent to which individuals believe that not compromising research integrity will damage one's career. Both models are estimated simultaneously using brms multivariate syntax, yielding joint posterior samples r for path *a* (prevalence -> mediator) and path *b* (mediator -> self-reported engagement). Minimally informative priors are specified for both mediation models, mirroring models 1-4. Again, we assign independent Cauchy prior log-odds, with scales 0.5 (for metric predictors) and 0.4 (for binary predictors). Demographic controls are included to increase precision. Prior predictive checks and sensitivity analyses are used to check for the influence of alternative prior distributions. Each model was fit to the 50 imputed datasets, and the posteriors were combined into a single inference. Each model was additionally fit to a complete-case dataset to determine robustness and sensitivity to missing data.

To test H5, testing whether researchers on average perceive other researchers to be engaging in QRPs more frequently than they themselves engage in them, we use a Bayesian alternative to the Wilcoxon signed-rank test for paired samples, using the Bayesian First Aid Alternative to the Wilcoxon Rank Test (Reinhardt, 2019). The test follows the estimation procedure outlined by Gelman et al. (2014), whereby the data are substituted by their ranks and transformed to the quantiles of a standard Gaussian distribution, using a Gaussian quantile function. This permits the modelling of the transformed data as a Gaussian distribution using Bayesian methods. Uniform priors are placed on these parameters (i.e., the mean and variance).

As with regression and mediation analysis, each model was fit to 50 imputed datasets, and the posteriors are combined into a single inference. We additionally estimated the model with a complete-case dataset to check sensitivity to missing data.

Following Kruschke (2014), we adopt the ROPE (95% HDI) index as an indicator of effect existence. We set our ROPE on the log-odds scale to ±0.18 because that range corresponds to what Cohen (1988) calls a "small" standardized effect (d ≈ 0.10) once you translate it into the logit (log-odds) metric. In other words, a very modest effect – one that would move someone from around 50% to about 55% probability – amounts to roughly 0.18 units on the logit scale, so any coefficient smaller than that is essentially negligible. This ROPE interval contains values that approximate the null value (0) and are suggestive of the absence of an effect. PD is indicative of the probability that the coefficient is strictly positive or negative, providing an estimate of the size of the effect (Makowski et al., 2019). For the paired-difference (Wilcoxon-style) test on the normal quantile scale, we use ±0.11 again reflecting a small effect (d ≈ 0.10) when data are on the standard-normal scale. Parameters are expressed as log-odds. Four chains were run for all models, at 5000 iterations (warm up = 1000) for model 1, 4000 iterations (warm up = 1000) for models 2 and 3 and 6000 iterations (warm up = 1000) for models 4-5. Iterations are adjusted because of varying levels of complexity across models. Inspection of trace plots and potential scale reduction factors (R-hat) indicate that all parameter estimates converged (R-hat < 1.01). All parameter estimates were stable for limits of all highest-density intervals (Effective Sample Size (ESS) > 5,000). Visual posterior predictive checks were undertaken for all models, and all models were deemed to adequately reproduce patterns observed in the real data, indicating good model fit. We additionally specified an adjacent category model for Model 4, providing estimates of category specific effects, allowing for the test of the proportional odds assumption.

Across all models MCMC convergence was assessed using R-hat (target <1.01), effective sample size and checked for divergent transitions. All models met these criteria are minor adjustment. Model fit is assessed using posterior predictive checks, which showed that simulated outcomes closely matched the observed data. To assess the proportional odds assumption in the ordinal specifications, we compared the cumulative logit models to an adjacent-category alternative. No meaningful violations were found, and substantiative inferences remained unchanged.

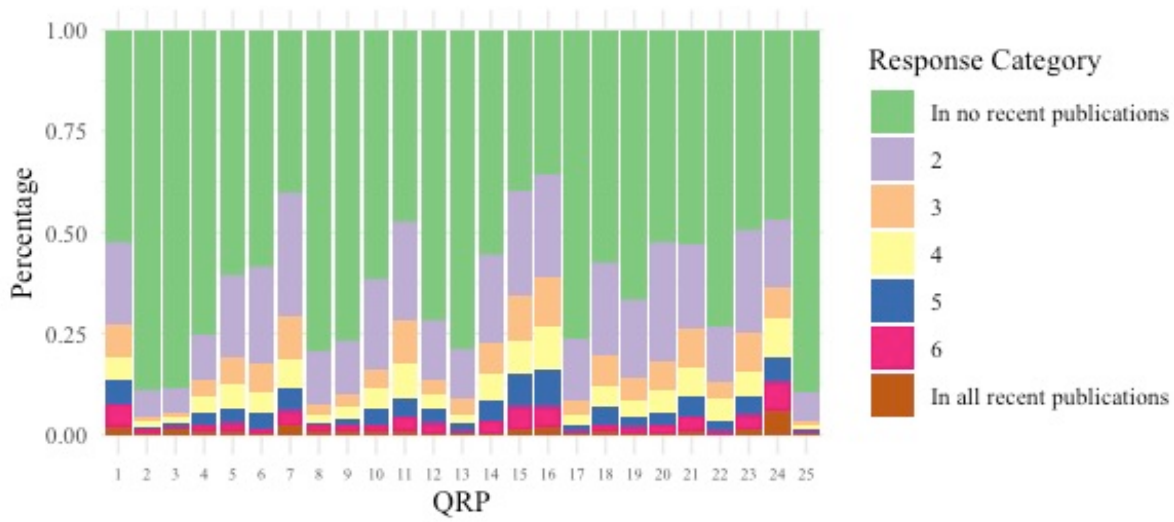
Results

Prevalence of QRPs

Figure 3.1 shows the percentage of observed answer categories for each of the twenty-five QRPs based on self-reported behaviour. The five most prevalent QRPs (i.e., QRPs with highest proportion of respondents scoring 1-7) are a) 'Deliberately citing one's own publications more than warranted by their relevance, to promote the visibility of one's work or improve one's citation metrics' (QRP 16: 64.5%), b) 'selectively citing irrelevant or unnecessary publications to please reviewers or editors' (QRP 15: 60.4%), c) 'citing literature deemed relevant for study without having actually read it' (QRP 7: 59.9%), and d) 'when reviewing a manuscript, not investing the effort necessary to conduct a thorough review' (QRP 11: 53.8%) and e) 'report main findings that turned out not to be "statistically significant", as evidence for no difference, or no association between study groups or variables' (QRP 24: 53.2%). The least prevalent QRPs were a) 'In a publication, failing to disclose relevant personal, financial, political or intellectual conflicts of interest' (QRP 3: 11.7%), b) 'failing to offer authorship to collaborators who had contributed sufficiently to the work to merit authorship' (QRP 2: 11.2%), and c) 'deliberately using another researcher's unpublished idea without giving credit' (QRP 25: 10.4%).

The QRPs perceived to be the most prevalent within respondent's fields (that is, practices with the greatest proportion of respondents scoring 2-7 on the Likert response scale) are a) deliberately [citing] own publications more than warranted by their relevance, to promote the visibility of your work or improve your citation metrics (QRP 16: 96.4%), b) citing literature deemed relevant for your study without actually having read it (QRP 7: 95.5%), c) without disclosure, formulate or change hypotheses after having seen the results, thereby presenting an unexpected finding as having been predicted from the start (QRP 20: 94.5%), and d) [reporting] main findings that turned out not to be "statistically significant", as evidence for no difference, or no association between study groups or variables' (QRP 23: 94.4%), and e) deliberately [dividing] the results of a study over more publications than needed with the intention to increase the number of publications (QRP 14: 94.3%). The QRPs judged to be least prevalent within respondent's fields were 'failing to offer authorship to collaborators who had contributed sufficiently to the work to merit authorship' (QRP 2: 74.4%) and 'in a publication, failing to disclose relevant personal, financial, political or intellectual conflicts of interest' (QRP 3: 70.5%).

Figure 3.1. Percentage of observed answer categories of QRPs across 4646 respondents.



The five QRPs perceived to be the most damaging to trust in research findings (i.e., QRPs with the highest proportion of respondents scoring 4, 5 and 6) are a) ‘Deliberately [refraining] from reporting findings that could weaken or contradict own theories, hypotheses or findings’ (QRP 19) b) ‘when reviewing a manuscript, not investing the effort necessary to conduct a thorough review’ (QRP 11), c) ‘selectively focusing on parts of the data or source material that support your preconceptions or hypotheses and deliberately disregarding parts that do not’ (QRP 18), d) ‘submitting a biased review report that evaluated the manuscript unfairly (QRP 12), and e) ‘avoiding to share data, research protocols, information on experimental setup, instrumentation, coding, or other information about a study requested by colleagues to evade transparency’ (QRP 9). ‘Deliberately [dividing] the results of a study over more publications than needed with the intention to increase the number of publications’ (QRP 14) and ‘deliberately [citing] own publications more than warranted by their relevance, to promote the visibility of your work or improve your citation metrics’ (QRP 16) were perceived to be the least damaging to scientific findings.

Ninety per cent of respondents reported to having engaged in at least one QRP more than never (i.e., Likert scale points 2-7), while 34.1% reported to having engaged in at least one QRP frequently (i.e., Likert scale points 5-7). The largest proportion of respondents reporting to have engaged in at least one QRP frequently were from the Natural and Technology Science (35.8%), followed by the Medical and Health Sciences (35.1%), the Social Sciences (31.3%) and then the Arts and Humanities (26%).

Regression Results

Table 3.3 shows summaries (incl. median posterior estimates, 95% higher density intervals (HDIs), probabilities of direction (PD) and percentages of 95% HDIs in region of practical equivalence (ROPE) for marginal posterior distributions of population and random effects for models 1-4. Table 3.3 shows that the effect of perceived extent of prevalence has a probability of 100% of being positive (Median = 1.95, 95%-CI [1.84, 2.05]) and is considered significant (0% of the HDI falls inside the ROPE). That is, the greater the extent to which respondents perceived QRPs to be prevalent within their field the greater the extent of their self-reported engagement in those practices. The greater the belief that a QRP is damaging to trust in scientific findings, the less likely a respondent is to engage in that practice (Median = -0.66, 95%-CI [-0.70, -0.63]). However, we do not find strong evidence for a moderating effect of belief that the QRP is damaging to trust in scientific findings on the relationship between perceptions of QRP prevalence and extent of self-reported engagement in QRPs (Median = -0.11, 95%-CI [-0.13, -0.09], *pd* = 100.00, ROPE (95%) = 100%). The random coefficient for variable 'perceived prevalence' is positive (Median = 0.74, 95%-CI [0.69, 0.78]), PD = 100%), indicating that on average, higher perceived prevalence predicts greater self-reported engagement. This random coefficient suggests that the measurement-level association between the extent to which a scientist perceives a QRP to be prevalent within their field and the extent to which they engage in that behaviour varies substantially between researchers. Additionally, the random coefficient for variable 'damage to trust in science findings' has a probability of 100.00 of being positive (Median = 0.54, 95%-CI [0.50, 0.59]), indicating that relationship between the extent to which a QRP is seen as damaging to trust in scientific findings and the extent of self-reported engagement in that practice varies between researchers.

In Model 4, respondent-level characteristics are introduced as covariates. In accordance with previous research, we find no evidence for gender differences in the extent of self-reported QRP engagement (Median = 0.15, 95%-CI [0.02, 0.27], *pd* = 99.99, ROPE (95%) = 100%). We find no disciplinary differences that are of practical relevance in the relationship between perceiving QRPs to be prevalent within own field and the extent of self-reported engagement in those practices (each ROPE (95%) > 2.5). Of substantiative interest is the unexpected finding of no relationship between research integrity training and extent of self-reported QRP engagement (Median = -0.1, 95%-CI [-0.14, 0.10], *pd* = 58.17, ROPE (95%) = 100.00). Moreover, disciplinary field is not considered to be a significant predictor of extent of self-reported QRP engagement. The relationship between career stage and extent of self-reported QRP engagement has a probability of 100% of being negative (Median = -0.63, 95%-CI [-0.76, -0.51]) and is significant (0% in ROPE (95%)). That is, a greater likelihood of engaging in QRPs is present for early-career researchers relative to late-career researchers. Cross-level interaction terms are additionally introduced in Model 4. We find no evidence that the relationships between perceptions of QRP prevalence and self-reported engagement is greater in the social sciences, relative to natural and technical sciences (Median = 0.01, 95%-CI [-0.05, 0.07], *pd* = 64.78, ROPE (95%) = 100.00), arts and humanities (Median = -0.17, 95%-CI [-0.24, -0.08], *pd* = 100.00, ROPE (95%) = 100.00) or the medical and health sciences (Median = 0.01, 95%-CI [-0.06, 0.08], *pd* = 65.05, ROPE (95%) = 100.00). Moreover, we find no evidence of a moderating association of 'career stage' on the relationship between perceptions of extent of QRP prevalence and extent of self-reported engagement in those practices (Median = -0.16, 95%-CI [-0.20, -0.05], *pd* = 100.00, ROPE (95%) = 100.00).

Table 3.4 shows a summary of model posterior estimates for the mediation analysis (Model 5), including Medians, 95% HDIs and PDs. Four chains were run at 4000 iterations (warm up = 1000). All parameter estimates were stable for limits of higher-density intervals (ESS > 10,000). Posterior predictive checks showed that the model sufficiently reproduced patterns observed in real data, suggesting good model fit. We do not find evidence for a mediating influence of perceptions of extent of QRP prevalence on extent of self-reported QRP engagement through the degree to which an individual believes not compromising research integrity is damaging to one's career (Median = 0.02, 95%-CI [0.02, 0.03], *pd* = 100.00, ROPE (95%) = 100.00). That is, the indirect association was found to be practically zero, with tight clustering of the posterior around zero.

Bayesian Wilcoxon Signed-Rank Paired Samples Test

As with the regression analyses, we adopt the ROPE (95% HDI) index as an indicator of effect existence. The ROPE is specified as ranging from -0.11 to 0.11, specified based on a negligible effect size (Cohen, 1988). PD is indicative of the probability that the coefficient is strictly positive or negative, providing an estimate of the size of the effect (Makowski et al., 2019). Three models were run at 2000 iterations each. Inspection of trace plots and R-hat indicate that all parameter estimates converged (R-hat < 1.01), while parameter estimates were stable for limits of HDIs (ESS > 3000). In support of Hypothesis 5, average estimates were higher when estimating the frequency of self-reported QRP engagement of others within their field (Median = 2.44), relative to making judgements about the frequency of one's own self-reported engagement in QRPs (Median = 0.55, Z = -1.29, 95%-CI [-1.31, -1.27], *pd* = 100.00) and this difference is considered significant (ROPE (95%) = 0.00).

Table 3.3: Summary of regression coefficients for cumulative ordinal models predicting extent of self-reported engagement in QRPs (PRINT data)

	Model 1 – Overall QRP Ordinal Regression Model			**Model 2 - Pooled Model Ordinal Regression Model**			**Model 3 - Partially Pooled Model (Level-1 Predictors) Ordinal Regression Model**			**Model 4 - Partially Pooled (Model with Interactions) Ordinal Regression Model**		
	Logit Median (95% Credible Interval)	PD (%)	ROPE (95%)	Logit Median (95% Credible Interval)	PD (%)	ROPE (95%)	Logit Median (95% Credible Interval)	PD (%)	ROPE (95%)	Median Logit (95% Credible Interval)	PD (%)	ROPE (95%)
Measurement Level Fixed Effects												
Perceived Prevalence				1.03 (1.01, 1.06)	100.00	0.00	1.60 (1.56, 1.65)	100.00	0.00	1.95 (1.84, 2.05)	100.00	0.00
Damage to Trust in Scientific Findings				-0.42 (-0.44, -0.40)	100.00	000	-0.72 (-0.75, -0.68)	100.00	0.00	-0.66 (-0.70, -0.63)	100.00	0.00
Prevalence * Damage to Trust in Scientific Findings										-0.11 (-0.13, -0.09)	100.00	100.00
Respondent Level Fixed Effects												
Overall Perceived Prevalence^	0.99 (0.93, 1.06)	100.00	0.00									
Overall Perceived Damage to Trust in Scientific Findings^	-0.29 (-0.34, -0.24)	100.00	96.58									
Gender (0 = Female)	0.08 (-0.06, 0.22)	87.89	100.00							0.15 (0.02, 0.27)	98.98	100.00
Country												
Denmark	0.29 (0.06, 0.54)	99.31	60.81							-0.07 (-0.23, 0.20)	56.07	100.00
Austria	0.45 (0.17, 0.76)	99.96	17.83							0.49 (0.23, 0.77(	99.99	9.27
Croatia	0.79 (0.47, 1.15)	100.00	0.00							0.53 (0.22, 0.84)	99.97	7.73
US	-0.17 (-0.62, 0.19)	80.91	77.20							-0.18 (-0.56, 0.18)	83.92	79.65
(0 = UK)												
Research Integrity Training (1 = Previously Undertaken)	0.09 (-0.04, 0.22)	90.29	100.00							-0.01 (-0.14, 0.10)	58.91	100.00
Career Stage (0 = Early Career)	-0.49 (-0.62, -0.36)	100.00	0.00							-0.63 (-0.76, -0.51)	100.00	0.00
Field												

Medical & Health	-0.06 (-0.24, 0.12)	72.99	100.00							0.15 (-0.02, 0.32)	95.71	100.00
Natural & Technical	0.15 (-0.02, 0.31)	96.17	100.00							0.03 (-0.12, 0.19)	67.20	100.00
Arts & Humanities	-0.04 (-0.25, 0.18)	64.59	100.00							-0.45 (-0.67, -0.25)	100.00	8.67
(0 = Social Sciences)												
Prevalence * Medical & Health										-0.05 (-0.15, 0.06)	81.00	100.00
Prevalence * Natural & Technical										0.12 (0.02, 0.22)	99.00	100.00
Prevalence * Arts & Humanities										0.05 (-0.09, 0.17)	75.00	100.00
Prevalence * Career Stage										-0.16 (-0.20, -0.05)	100.00	100.00
Intercept												
Random Effects												
Variance (Perceived Prevalence)							0.76 (0.72, 0.81)			0.74 (0.69, 0.78)		
Variance (Damage to Scientific Findings)							0.56 (0.51, 0.60)			0.54 (0.50, 0.59)		
Variance (Intercept)							1.91 (1.85, 1.97)			1.85 (1.79, 1.91)		
Covariance (Prevalence * Intercept)							0.07 (0.01, 0.13)			0.11 (0.05, 0.17)		
Covariance (Damage to Scientific Findings * Intercept)							0.24 (0.17, 0.31)			0.18 (0.11, 0.25)		
Covariance (Prevalence * Damage to Scientific Findings)							0.01 (-0.08, 0.10)			0.08 (-0.01, 0.18)		

^ Overall QRP and Overall Prevalence computed as summative score and recoded into set of ordered categories; % in ROPE indicates % of 95% HDI posterior in ROPE;

Table 3.4: Mediation analysis of the mechanism of 'compromising behaviour out of career necessity' on the relationship between perceived QRP prevalence and self-reported QRP self-reported engagement.

	Median Logit (95% Credible Interval)	PD (%)	ROPE (95%)
Model 5A **Outcome: Compromising Research Integrity out of Career Necessity**			
Intercept			
Prevalence (Overall)	0.23 (0.18, 0.27)	100.00	0.00
Gender (1 = Male)	0.05 (-0.05, 0.16)	83.89	100.00
Career Stage (1 = Early Career)	-0.24 (-0.35, -0.14)	100.00	96.24
Research Integrity Training	-0.02 (-0.13, 0.08)	67.97	100.00
Country			
Austria	0.24 (0.02, 0.47)	93.34	78.66
Denmark	0.19 (-0.32, 0.17)	98.28	95.00
Croatia	-0.07 (-0.32, 0.17)	72.15	100.00
United States	-0.53 (-0.91, -0.16)	99.74	13.13
(0 = UK)			
Field			
Arts & Humanities	-0.01 (-0.19, 0.16)	55.82	100.00
Nature & Technical	0.08 (-0.06, 0.22)	86.37	100.00
Medical & Health	0.01 (-0.14, 0.16)	54.36	100.00
(0 = Social Science)			
Model 5B **Outcome: Overall Extent of self-reported Engagement in QRPs**			
Intercept			
Prevalence (Overall)	0.86 (0.80, 0.92)	100.00	0.00
Compromising Research Integrity out of Career Necessity	0.10 (0.07, 0.14)	100.00	100.00
Gender (1 = Male)	0.17 (0.04, 0.30)	99.50	100.00
Career Stage (1 = Early Career)	-0.45 (-0.58, -0.32)	100.00	1.05
Research Integrity Training	0.06 (-0.06, 0.19)	83.98	100.00
Austria	0.34 (0.06, 0.63)	99.28	46.88
Denmark	0.18 (-0.05, 0.42)	100.00	5.20
Croatia	0.58 (0.25, 0.92)	93.37	91.91
United States	-0.11 (-0.52, 0.27)	70.53	88.56
(0 = UK)			
Field			
Arts & Humanities	-0.05 (-0.27, 0.17)	68.58	100.00
Nature & Technical	0.14 (-0.03, 0.31)	94.55	100.00
Medical & Health	-0.08 (-0.26, 0.11)	80.43	100.00
(0 = Social Science)			
Indirect Association			
	0.02 (0.02, 0.03)	100.00	100.00

Discussion

The widespread prevalence of QRPs across disciplinary fields is now well evidenced. This prevalence, combined with the fact that QRPs have been found to undermine the replicability and reliability of scientific findings, suggest a need to identify the factors and mechanisms that facilitate self-reported engagement in these behaviours. In recognising the correlates of self-reported QRP engagement, interventional strategies can be developed that seek to mitigate their use (Begley & Ioannidis, 2015; Fanelli, 2009; John et al., 2012). With this objective in mind, the current study sought to examine the extent to which descriptive norms correlate with researchers' research conduct – specifically, the extent to which researchers engage in QRPs. To achieve this, we utilised data derived from a cross-national survey administered to a representative sample of researchers operating across disciplinary fields. While past studies have conceptualised QRPs as a mean score indicating average self-reported engagement across all QRPs, we used a novel multilevel modelling approach to assess the 'measurement-level' relationship between perceptions of prevalence and extent of self-reported engagement in QRPs. In doing so, we made no distinction between the qualitative nature of each the practices in this study. This allowed us to undertake a robust and valid assessment of the relationship between the extent to which QRPs are perceived as prevalent and the degree of self-reported engagement in those practices.

Main disciplinary field was *not* associated with self-reported QRP engagement. This suggests that, in the present sample, QRP engagement did not vary systematically across broad disciplinary categories. However, this null finding should also be interpreted cautiously. Disciplinary field is likely to capture heterogeneous differences in research methods, publication norms, collaborative structures, measurement practices, and exposure to competitive or productivity-based incentives. Moreover, specific QRPs may not be equally applicable, recognisable, or normatively defined across disciplines. Practices such as selective reporting, HARKING, excluding data, or failing to report null findings may have clearer meanings in experimental, quantitative, or hypothesis-testing traditions than in interpretive, theoretical, historical, or qualitative research. Thus, the absence of an association should not be taken to imply that disciplinary context is irrelevant to research integrity. Rather, broad disciplinary categories may be insufficiently precise to capture the specific normative, methodological, and institutional contexts that shape opportunities for, and interpretations of, QRP engagement.

In support of our primary hypothesis (H1), we found that the degree to which researchers perceived QRPs to be prevalent within their disciplinary field, the more likely they were to have engaged in those practices to a greater extent. This aligns with research findings on other topics which suggest that descriptive norms – indicating which behaviours are typical within a particular context – can undermine injunctive norms fostering ethical behaviour (Cialdini et al., 1990; Staunton et al., 2014). Our findings build on those from Sacco et al. (2018) study of National Institutes of Health principal investigators, which found that the more individuals perceived QRPs to be normative, the more ethically defensible they reported them to be and the more willing they were to engage in them. Here, we go further by demonstrating that a direct relationship exists between descriptive norms and the *actual*, albeit self-reported, conduct of researchers. Our findings are consistent with the findings from Rajah-Kanagasabai and Roberts' (2015) study that identified a relationship between perceptions of descriptive norms and research behaviours among a sample of university students in Australia. Extending this, our research demonstrates similar effects among academic faculty across disciplines and countries.

We hypothesised that this relationship is mediated through a belief that not compromising one's research integrity will have a negative association with one's career (H2). Using neutralisation theory as a framework, Sacco et al. (2018) argued that if scientists perceive others to be using QRPs to their advantage, their non-engagement in QRPs puts them at a competitive disadvantage, which in turns motivates engagement in those practices. This 'defence of necessity' argument is logical. Past research has demonstrated a relationship between competition and engagement in QRPs (Gopalakrishna et al., 2022). Moreover, the social psychological literature shows that perceptions of competition incentivise individuals to improve their position relative to their competitors, such as through cheating or engaging in unethical behaviour (Belot & Scröder, 2013). Perceiving engagement in QRPs to be typical likely contributes to this perception, thus motivating researchers to engage in those practices in order remain competitive. However, we failed to empirically validate such a mechanism. This is contrary to Sacco et al.'s (2018) common-sense deductions that the statistically significant relationship between perceptions of normativity and willingness to engage in QRPs is a result of these perceptions contributing to a belief that it is necessary to engage in these behaviours to keep up with colleagues.

Therefore, it is feasible that this relationship is the product of other social-psychological mechanisms. Researchers may base their beliefs about the appropriateness and justifiability of specific research practices on what they believe is typical within their disciplinary field. That is, these descriptive norms may influence research behaviour through a process of 'informational social influence', whereby individuals conform to the behaviours of others on the basis that a) the wide use of the QRP points to the appropriateness of the practice (Festinger et al., 1950; Hardin & Higgins, 1996; Turner, 1991) and b) the individual has a desire to be correct (Jones & Gerrard, 1967). This process typically occurs when decision-making is characterised by ambiguity (Festinger, 1954; Klucharev et al., 2009; Toelch & Dolan, 2015). Certainly, given that *some* QRPs may be deemed appropriate in certain contexts but not in others, their use may be sometimes considered ambiguous, thus motivating researchers to look to others for guidance on whether it is appropriate to engage. Relatedly, the influence of descriptive norms on researcher's behaviour may in part occur through the process of 'normative social influence' (Asch, 1955, 1956; Asch & Guetzkow, 1951), whereby individuals conform to behaviours that they perceive as 'typical' because of their need to feel as if they belong to their scholarly or disciplinary communities.

In line with the dominant thesis that systemic pressures drive engagement in QRPs, early-career researchers are often assumed to be more susceptible to cutting corners, as such pressures are particularly acute at this stage of the academic career. In this study we found no empirical evidence to support this. This indicates that such mechanisms of influence are potentially equally effectual regardless of career stage. On the one hand, early-career researchers may be more vigilant of, and motivated to identify with, the most 'appropriate' research practices, and look at the most common behaviours as a basis of justifiability. However, mid-late career researchers are likely to have *already* aligned their behaviours to that of the community. That is, all researchers appear to align their behaviours with the typical behaviour in their respective communities, but they are in different stages of the alignment process. Moreover, such influence seems to be unwavering across disciplinary fields. We found no disciplinary differences in the relationship between perceptions of prevalence and behaviour. Such a finding is theoretically sensible. Wherever QRPs are relevant to research practice and seen as typical, researchers are equally likely to engage, regardless of the field in which they operate.

Contrary to our predictions, the extent to which the behaviours were seen as damaging to trust in scientific findings did not correlate with the strength of the relationship between perceptions of QRP prevalence and extent of self-reported engagement in those practices. We expected descriptive norms to command the most influence on research practice when the behaviours were perceived as more damaging to trust in scientific findings. Where the behaviours were perceived as undamaging, we expected researchers to engage regardless of the behaviour of others. That is, we didn't expect to identify a relationship between estimates of prevalence and extent of engagement for less damaging behaviours, because we expected this engagement to be driven primarily by utility in these instances. However, our findings suggest that researchers look to the behaviour of others as a guide for their own behaviour even when those practices are perceived as undamaging to science. This suggests that an individual's desire to conform may transcend considerations of ethicality.

Regardless of the mechanisms through which descriptive norms influence research practice, in empirically validating such a relationship, this study motivates the consideration of novel interventions that can be used to reduce self-reported engagement in QRPs. For example, interventional campaigns should seek to counter the inference that because a QRP is widely used, it's use is justified. However, given that descriptive norms were related to extent of self-reported QRP engagement, *regardless* of whether the practice was seen as damaging to trust in scientific findings, manipulating 'injunctive norms' – that is, pointing towards the unethical nature of the practice – may not be the most efficacious means of mitigating engagement. The findings allude to a caution that should be taken when emphasising the descriptive normativity of QRPs, including those disseminated through institutional messages. In contexts that are resistant to imposition of authoritative enforcement, such as scholarly communities, such interventions may be more efficacious and 'accepted'. Certainly, such interventions sit more comfortably in a system that is self-correcting and based on mutual trust between actors. Similar interventions have been used widely in social norms marketing, whereby information about a groups behaviour or opinions is made known using mass communication materials (e.g., Perkins & Craig, 2006). There is the potential for such interventions to be effective in mitigating the use of QRPs.

Limitations

While the study makes meaningful theoretical contributions to the field, it is necessary to consider its limitations. First, as the study adopted a cross-sectional design it is not possible to establish the causal direction of the relationship between perceptions of prevalence and extent of self-reported engagement in QRPs. Although the findings are consistent with the possibility that perceptions of widespread QRP engagement may be linked to researchers' own reported engagement, alternative interpretations are equally plausible.

One possibility is reverse causality. Researchers who report engaging in QRPs may infer that such practices are also common among their peers, extrapolating from their own behaviour when estimating the prevalence of QRPs within their field. This interpretation is consistent with research on the false-consensus effect, whereby individuals tend to overestimate the extent to which others share their own behaviours, beliefs, or attitudes (Ross et al., 1977; Sherman et al., 1983; van der Pligt, 1984; Mullen et al., 1985; Marks & Miller, 1987). In this context, self-reported QRP engagement may therefore shape estimates of QRP prevalence, rather than perceived prevalence shaping engagement.

A further possibility is that the observed association is non-causal. Perceived QRP prevalence and self-reported QRP engagement may both reflect broader features of the research environment, such as disciplinary norms, publication pressure, perceived competition, institutional expectations, or local research cultures. Researchers working in environments perceived as highly competitive or permissive of questionable practices may both estimate higher levels of QRP prevalence and report greater personal engagement in such practices, without either variable directly causing the other. Similarly, individual differences in cynicism, norm perception, willingness to disclose, or interpretation of the QRP items may contribute to the observed association.

It is also important to emphasise that the outcome measure captures reported engagement in QRPs, rather than directly observed behaviour. Although self-report measures are commonly used in research on sensitive behaviours, responses may be affected by recall error, social desirability concerns, differences in interpretation of what constitutes a QRP, and variation in willingness to disclose questionable practices. Accordingly, the findings should be interpreted as associations involving self-reported QRP engagement, rather than as direct evidence of actual behavioural engagement.

Taken together, these limitations mean that the results should not be interpreted as demonstrating that perceptions of QRP prevalence cause researchers to engage in QRPs. Rather, they indicate that perceived prevalence and self-reported engagement are associated in theoretically meaningful ways. Future research using longitudinal, experimental, or behavioural designs would be needed to assess whether perceived prevalence precedes changes in QRP engagement, whether self-reported engagement shapes perceived prevalence, or whether both are jointly influenced by broader individual, institutional, or disciplinary factors.

Moreover, the use of self-report measures is likely to result in the under-reporting of behaviour. Certainly, John et al. (2012) using scoring algorithms and incentives for truth-telling, found that QRP prevalence rates without incentives are highly conservative. However, research has demonstrated that self-reports on dishonest behaviour and true rates of engagement are often similar (Halevy et al., 2014). Thus, we can be confident in using self-report measures as proxies for actual behaviour, albeit such estimates will be conservative.

A related limitation concerns the possibility of non-response bias arising from low response rates. This concern is salient given the sensitive nature of self-reported QRP engagement. On this basis, participation in the survey may have been non-random, whereby researchers with a strong interest in research integrity, those more willing to reflect or disclose engagement in QRPs, and those with particular experiences with systemic pressures may be more likely to respond. On the contrary, those who had previously engaged in QRPs, or perceived such disclosures as risky, may have been less inclined to participate. On this basis, the generated sample may differ systematically from the wider population of academics and researchers.

This selectivity has implications for both descriptive and inferential interpretation. Prevalence estimates should not be treated as precise estimates, since they may be affected by social desirability bias and differential participation. Moreover, selective participation may have restricted the observed range of both the outcome and key predictors. For instance, if respondents were disproportionately similar in their concern about research integrity, willingness to disclose, or exposure to research-related pressures, this may have reduced variance in self-reported QRP engagement and attenuated associations with predictor variables. On this basis, the results are best interpreted as evidence of associations within the responding sample, rather than as definitive estimates of population-level prevalence or effect sizes.

Effect sizes across predictors are uniformly small. This is consistent with patterns observed in the extant literature, where correlates of QRP engagement tend to be modest but potentially meaningful. Certainly, self-reported engagement is likely to conservatively estimated, and the restricted variance in QRP engagement – as well as predictors - mean that associations are likely to be reduced in magnitude. That is, small effect sizes are likely to reflect restriction I range rather than the absence of substantively meaningful relationships. Furthermore, small shifts in the probability of self-reported QRP engagement may translate into meaningful differences when aggregated at across large numbers of researcher, studies and publications. Thereby, modest predictor effects may be practically significant. However, the modest magnitudes reinforce the need for humility in interpretation, particularly considering possible attenuation and nonresponse-related biases.

Finally, while a particular strength of this study is the representativeness of the sample, with researchers from a diversity of disciplinary fields and at various stages of their career, it focuses wholly on western European and Anglosphere countries. On this basis, findings are not generalisable beyond these unique cultural settings. We recommend that future research seeks to make cross-cultural comparisons, allowing for stronger tests of the hypotheses.

Conclusion

In this research we sought to recognise the role that perceptions of QRP prevalence play in facilitating engagement in QRPs. We found that descriptive norms were predictive of extent of self-reported QRP engagement. This relationship seems to be found across disciplinary fields and career stages. The relationship is additionally unrelated to beliefs about the extent to which QRPs are seen as damaging to trust in scientific findings. We failed to find a mediating role of self-reported engagement out of career necessity in the relationship between perceptions of prevalence and behaviour, and other mechanisms might facilitate the influence of descriptive norms. Future research should seek to elucidate other mechanisms that may facilitate this relationship, such as through informational social influence. In doing so, we open the door for the development of novel interventions to tackle QRP engagement, such as those grounded in social psychology and behavioural economics.

References

Anderson, L. R., Mellor, J. M., & Milyo, J. (2004). Social Capital and Contributions in a Public-Goods Experiment. *The American Economic Review*, *94*(2), 373–376. http://doi.org/10.1257/0002828041302082

Anderson, M. S., Horn, A. S., Risbey, K. R., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a National Survey of NIH-funded scientists. *Academic medicine : journal of the Association of American Medical Colleges*, *82*(9), 853–860. https://doi.org/10.1097/ACM.0b013e31812f764c

Antes, A. L., Wang, X., Mumford, M. D., Brown, R. P., Connelly, S., & Devenport, L. D. (2010). Evaluating the effects that existing instruction on responsible conduct of research has on ethical decision making. *Academic medicine : journal of the Association of American Medical Colleges*, *85*(3), 519–526. https://doi.org/10.1097/ACM.0b013e3181cd1cc5

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men; research in human relations* (pp. 177–190). Carnegie Press.

Asch, S. E. (1955). Opinions and social pressure. *Scientific American, 193*(5), 31–35. https://doi.org/10.1038/scientificamerican1155-31

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied, 70*(9), 1–70. https://doi.org/10.1037/h0093718

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied, 70*(9), 1–70. https://doi.org/10.1037/h0093718

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

Begley, C. G., & Ioannidis, J. P. (2015). Reproducibility in science: improving the standard for basic and preclinical research. *Circulation research*, *116*(1), 116–126. https://doi.org/10.1161/CIRCRESAHA.114.303819

Belot, M., & Schröder, M. (2013). Sloppy work, lies and theft: A novel experimental design to study counterproductive behaviour. *Journal of Economic Behavior & Organization*, *93*, 233-238. https://doi.org/10.1016/j.jebo.2013.03.019

Belot, M., & Schröder, M. (2013). Sloppy work, lies and theft: A novel experimental design to study counterproductive behaviour. *Journal of Economic Behavior & Organization*, *93*, 233-238. https://doi.org/10.1016/j.jebo.2013.03.019

Berger, J., & Berliner, L. M. (1986). Robust Bayes and Empirical Bayes Analysis with ε-Contaminated Priors. *The Annals of Statistics*, *14*(2), 461–486. http://doi.org/10.1214/aos/1176349933

Bertholet, N., Daeppen, J. B., Cunningham, J. A., Burnand, B., Gmel, G., & Gaume, J. (2016). Are young men who overestimate drinking by others more likely to respond to an electronic normative feedback brief intervention for unhealthy alcohol use?. *Addictive behaviors*, *63*, 97–101. https://doi.org/10.1016/j.addbeh.2016.07.015

Bicchieri, C. (2017). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York, online ed, Oxford Academic. https://doi.org/10.1093/acprof:oso/9780190622046.001.0001, Brooker, R., & Allum, N. (2024). Investigating the links between questionable research practices, scientific norms and organisational culture. *Research integrity and peer review*, *9*(1), 12. https://doi.org/10.1186/s41073-024-00151-x

Bürkner, P-C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1-28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science, 2*(1), 77–101. https://doi.org/10.1177/2515245918823199

Butterfield, K. D., Treviño, L. K., & Weaver, G. R. (2000). Moral awareness in business organizations: Influences of issue-related and social context factors. *Human Relations, 53*(7), 981–1018. https://doi.org/10.1177/0018726700537004

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science (New York, N.Y.)*, *351*(6280), 1433–1436. https://doi.org/10.1126/science.aaf0918

Camerer, C., Dreber, A., Holzmeister, F., Ho, T., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B.A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., … Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour, 2*, 637 - 644. https://doi.org/10.1038/s41562-018-0399-z

Chapman, K. J., Lupton, R. A. (2004). Academic dishonesty in a global educational market: a comparison of Hong Kong and American university business students. *International*

Journal of Educational Management, *18*(7), 425–435.
https://doi.org/10.1108/09513540410563130

Chavalarias, D., Wallach, J. D., Li, A. H., & Ioannidis, J. P. (2016). Evolution of Reporting P Values in the Biomedical Literature, 1990-2015. *JAMA*, *315*(11), 1141–1148. https://doi.org/10.1001/jama.2016.1952

Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K., & Winter, P. L. (2006). Managing social norms for persuasive impact . *Social Influence*, *1*(1), 3–15. https://doi.org/10.1080/15534510500181459

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. *Advances in Experimental Social Psychology*, *24*, 201-234. https://doi.org/10.1016/S0065-2601(08)60330-5

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd. ed.). Lawrence Erlbaum Associates, Inc.

Dwan, K., Gamble, C., Williamson, P. R., Kirkham, J. J., & Reporting Bias Group (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PloS one*, *8*(7), e66844. https://doi.org/10.1371/journal.pone.0066844

Fanelli D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PloS one, 4(5), e5738. https://doi.org/10.1371/journal.pone.0005738

Fanelli, D. (2018) 'Is science really facing a reproducibility crisis, and do we need it to?', *Proceedings of the National Academy of Sciences*, 115(11), pp. 2628–2631. Available at: https://doi.org/10.1073/pnas.1708272114 (Accessed: 3 June 2026).

Fanelli, D., Costas, R., & Larivière, V. (2015). Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. *PloS one*, *10*(6), e0127556. https://doi.org/10.1371/journal.pone.0127556

Fang, F. C., Bennett, J. W., & Casadevall, A. (2013). Males are overrepresented among life science researchers committing scientific misconduct. *mBio*, *4*(1), e00640-12. https://doi.org/10.1128/mBio.00640-12

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2)*,* 117–140. https://doi.org/10.1177/001872675400700202

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2)*,* 117–140. https://doi.org/10.1177/001872675400700202

Festinger, L., Schachter, S., & Back, K. (1950). *Social pressures in informal groups; a study of human factors in housing.* Harper.

Fiedler, K., & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological and Personality Science, 7*(1), 45–52. https://doi.org/10.1177/1948550615612150

Fiedler, K., & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological and Personality Science, 7*(1), 45–52. https://doi.org/10.1177/1948550615612150

Francis, G. (2012) 'Publication bias and the failure of replication in experimental psychology', *Psychonomic Bulletin & Review*, 19(6), pp. 975–991. Available at: https://doi.org/10.3758/s13423-012-0322-y (Accessed: 3 June 2026).

Franco, A., Malhotra, N., & Simonovits, G. (2014). Social science. Publication bias in the social sciences: unlocking the file drawer. *Science (New York, N.Y.)*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models. *The Annals of Applied Statistics*, *2*(4), 1360–1383. https://doi.org/10.1214/08-AOAS191

Gerber, A. S., & Rogers, T. (2009). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics, 71*(1), 178–191. https://doi.org/10.1017/S0022381608090117

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science". *Science (New York, N.Y.)*, *351*(6277), 1037. https://doi.org/10.1126/science.aad7243

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35*(3), 472–482. https://doi.org/10.1086/586910

Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PloS one*, *17*(2), e0263023. https://doi.org/10.1371/journal.pone.0263023

Halevy, R., Shalvi, S., & Verschuere, B. (2014). Being honest about dishonesty: Correlating self-reports and actual lying. *Human Communication Research, 40*(1), 54–72. https://doi.org/10.1111/hcre.12019

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition, Vol. 3. The interpersonal context* (pp. 28–84). The Guilford Press.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS biology*, *13*(3), e1002106. https://doi.org/10.1371/journal.pbio.1002106

Holm, S., & Hofmann, B. (2018). Associations between attitudes towards scientific misconduct and self-reported behavior. *Accountability in research*, *25*(5), 290–300. https://doi.org/10.1080/08989621.2018.1485493

Ioannidis J. P. (2005). Why most published research findings are false. *PLoS medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

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. https://doi.org/10.1177/0956797611430953

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. https://doi.org/10.1177/0956797611430953

Johnson V. E. (2013). Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(48), 19313–19317. https://doi.org/10.1073/pnas.1313476110

Johnson, L., Erikson, S., Helgesson, G., & Hansson, M. G. (2014). Making researchers moral: Why trustworthiness requires more than ethics guidelines and review. *Research Ethics*, *10*(1), 29-46. https://doi.org/10.1177/1747016113504778

Jones, E.E., & Gerard, H. (1967). *Foundations of social psychology.* Wiley.

Kerr N. L. (1998). HARKing: hypothesizing after the results are known. *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc*, *2*(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., & Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, *61*(1), 140–151. https://doi.org/10.1016/j.neuron.2008.11.027

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. (2nd Ed.). Academic Press.

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review, 25*(1), 178–206. https://doi.org/10.3758/s13423-016-1221-4

Lange, K. L., Little, R. J. A., & Taylor, J. M. G. (1989). Robust Statistical Modeling Using the *t* Distribution. *Journal of the American Statistical Association*, *84*(408), 881–896. https://doi.org/10.1080/01621459.1989.10478852

Larson, R. C., Ghaffarzadegan, N., & Xue, Y. (2014). Too Many PhD Graduates or Too Few Academic Job Openings: The Basic Reproductive Number R_0 in Academia. *Systems research and behavioral science*, *31*(6), 745–750. https://doi.org/10.1002/sres.2210

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of Effect Existence and Significance in the Bayesian Framework. *Frontiers in psychology*, *10*, 2767. https://doi.org/10.3389/fpsyg.2019.02767

Marks, G., & Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin, 102*(1), 72–90. https://doi.org/10.1037/0033-2909.102.1.72

Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review, 22,* 635–659. https://doi.org/10.2307/2089193

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS biology*, *16*(3), e2004089. https://doi.org/10.1371/journal.pbio.2004089

Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E., & Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology, 21*(3), 262–283. https://doi.org/10.1016/0022-1031(85)90020-4

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature human behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Oberauer, K. and Lewandowsky, S. (2019) 'Addressing the theory crisis in psychology', *Psychonomic Bulletin & Review*, 26(5), pp. 1596–1618. Available at: https://doi.org/10.3758/s13423-019-01645-2 (Accessed: 3 June 2026).

Open Science Collaboration. (2015). PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science (New York, N.Y.)*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Paruzel-Czachura, M., Baran, L., & Spendel, Z. (2020). Publish or be ethical? Publishing pressure and scientific misconduct in research. *Research Ethics*, *17*(3), 375-397. https://doi.org/10.1177/1747016120980562

Perkins, H. W., & Craig, D. W. (2006). A Successful Social Norms Campaign to Reduce Alcohol Misuse Among College Student-Athletes. *Journal of Studies on Alcohol, 67*(6), 880–889. https://doi.org/10.15288/jsa.2006.67.880

Perkins, H. W., Craig, D. W., & Perkins, J. M. (2011). Using social norms to reduce bullying: A research intervention among adolescents in five middle schools. *Group Processes & Intergroup Relations, 14*(5), 703–722. https://doi.org/10.1177/1368430210398004

Rajah-Kanagasabai, C. J., & Roberts, L. D. (2015). Predicting self-reported research misconduct and questionable research practices in university students using an augmented Theory of Planned Behavior. *Frontiers in psychology*, *6*, 535. https://doi.org/10.3389/fpsyg.2015.00535

Rajah-Kanagasabai, C. J., & Roberts, L. D. (2015). Predicting self-reported research misconduct and questionable research practices in university students using an augmented Theory of Planned Behavior. *Frontiers in psychology*, *6*, 535. https://doi.org/10.3389/fpsyg.2015.00535

Reinhardt, J. (2019, 10 May). *bayes.wilcox.test: Bayesian First Aid Alternative to the Wilcoxon Rank Test*. Rdrr. https://rdrr.io/github/joereinhardt/BayesianFirstAid-Wilcoxon/man/bayes.wilcox.test.html

Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13*(3), 279–301. https://doi.org/10.1016/0022-1031(77)90049-X

Rubin, M. (2023) 'Questionable metascience practices', *Journal of Trial & Error*, 4(1), pp. 5–20. Available at: https://doi.org/10.36850/mr4 (Accessed: 3 June 2026).

Rubin, M. (2025) *A brief review of research that questions the impact of questionable research practices*. PsyArXiv. Available at: https://doi.org/10.31234/osf.io/ah9wb_v3 (Accessed: 3 June 2026).

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In Defense of the Questionable: Defining the Basis of Research Scientists' Engagement in Questionable Research Practices. *Journal of empirical research on human research ethics : JERHRE*, *13*(1), 101–110. https://doi.org/10.1177/1556264617743834

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In Defense of the Questionable: Defining the Basis of Research Scientists' Engagement in Questionable Research Practices. *Journal of empirical research on human research ethics : JERHRE*, *13*(1), 101–110. https://doi.org/10.1177/1556264617743834

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In Defense of the Questionable: Defining the Basis of Research Scientists' Engagement in Questionable Research Practices. *Journal of empirical research on human research ethics : JERHRE*, *13*(1), 101–110. https://doi.org/10.1177/1556264617743834

Schneider, J. W. (2018). PRINT – Practices, Perceptions, and Patterns of Research Integrity. OSF; 2018. Available from https://osf.io/rf4bn/.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological science*, *18*(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Sherif, M. (1936). The psychology of social norms. Harper.

Sherman, S. J., Presson, C. C., Chassin, L., Corty, E., & Olshavsky, R. (1983). The false consensus effects in estimates of smoking prevalence: Underlying mechanisms. *Personality and Social Psychology Bulletin, 9*(2), 197–207. https://doi.org/10.1177/0146167283092003

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. https://doi.org/10.1177/0956797611417632

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. https://doi.org/10.1177/0956797611417632

Smyth, M. L., & Davis, J. R. (2004). Perceptions of Dishonesty among Two-Year College Students: Academic versus Business Situations. *Journal of Business Ethics*, *51*(1), 63–73. https://doi.org/10.1023/B:BUSI.0000032347.79241.3c

Stangor, C., Sechrist, G. B., & Jost, J. T. (2001). Changing racial beliefs by providing consensus information. *Personality and Social Psychology Bulletin, 27*(4), 486–496. https://doi.org/10.1177/0146167201274009

Staunton, M., Louis, W. R., Smith, J. R., Terry, D. J., & McDonald, R. I. (2014). How negative descriptive norms for healthy eating undermine the effects of positive injunctive norms. *Journal of Applied Social Psychology, 44*(4), 319–330. https://doi.org/10.1111/jasp.12223

Staunton, M., Louis, W. R., Smith, J. R., Terry, D. J., & McDonald, R. I. (2014). How negative descriptive norms for healthy eating undermine the effects of positive injunctive norms. *Journal of Applied Social Psychology, 44*(4), 319–330. https://doi.org/10.1111/jasp.12223

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American Statistical Association, 54,* 30–34. https://doi.org/10.2307/2282137

Stone, T. H., Jawahar, I. M., & Kisamore, J. L. (2009). Using the theory of planned behavior and cheating justifications to predict academic misconduct. *The Career Development International, 14*(3), 221–241. https://doi.org/10.1108/13620430910966415

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

Tankard, M. E., & Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review, 10*(1), 181–211. https://doi.org/10.1111/sipr.12022

Thomas, W. I., & Thomas, D. S. (1928). The child in America. Knopf.

Tijdink, J. K., Verbeke, R., & Smulders, Y. M. (2014). Publication pressure and scientific misconduct in medical scientists. *Journal of empirical research on human research ethics : JERHRE*, *9*(5), 64–71. https://doi.org/10.1177/1556264614552421

To, C., Kilduff, G. J., & Rosikiewicz, B. L. (2020). When Interpersonal Competition Helps and When It Harms: An Integration via Challenge and Threat. *The Academy of Management Annals*, *14*(2), 908-934. https://doi.org/10.5465/annals.2016.0145

Toelch, U., & Dolan, R. J. (2015). Informational and Normative Influences in Conformity from a Neurocomputational Perspective. *Trends in cognitive sciences*, *19*(10), 579–589. https://doi.org/10.1016/j.tics.2015.07.007

Turner, J. C. (1991). *Social influence.* Thomson Brooks/Cole Publishing Co.

Ulrich, R. and Miller, J. (2020) 'Questionable research practices may have little effect on replicability', *eLife*, 9, e58237. Available at: https://doi.org/10.7554/eLife.58237 (Accessed: 3 June 2026).

van Buuren, S., & Groothuis-Oudshoorn, C. G. M. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of statistical software*, *45*(3). http://www.jstatsoft.org/v45/i03

Van der Pligt, J. (1984). Attributions, false consensus, and valence: Two field studies. *Journal of Personality and Social Psychology, 46*(1), 57–68. https://doi.org/10.1037/0022-3514.46.1.57

Xie, Y., Wang, K., & Kong, Y. (2021). Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis. *Science and engineering ethics*, *27*(4), 41. https://doi.org/10.1007/s11948-021-00314-9

Chapter 4

Motivations, Pressures, and Personality: What Drives Questionable Research Practices?

Abstract

This study examines the social and psychological correlates of questionable research practices (QRPs), including personality traits, commitment to Mertonian scientific norms, personal and proximal science-specific motivations, political orientation, epistemic orientation, and pressures related to publication and funding. Using data from a large-N cross-sectional survey of researchers operating across disciplinary fields in the European and Anglo-American spheres, we find that commitment to universalism, communalism, and organised scepticism – as well as higher conscientiousness, agreeableness, neuroticism, and openness – are negatively associated with self-reported QRP engagement. Publication pressure predicts greater self-reported QRP engagement, but this effect is attenuated by stronger endorsement of universalism, suggesting that commitment to Mertonian norms may mitigate the influence of systemic pressures. Funding pressure was not statistically associated with QRPs. Motivations to falsify theories and findings, and to improve impact and quality of results, are negatively associated with self-reported QRP engagement. Conversely, career-achievement motivations (e.g., increased salary, promotion etc.) and esteem-related motivations (e.g., fame, achieving social and political goals) are positively associated with self-reported QRP engagement, partly via their relationship with scientific achievement- and impact-driven goals. These findings highlight specific social and psychological factors that correlate with self-reported QRP engagement, offering potential leverage points for efforts to promote research integrity and reduce questionable practices.

Key words: Questionable Research Practices (QRPs), Research Integrity, Mertonian Norms, Publication Pressure, Personality Traits, Motivations

Competing Interests: Robin Brooker is supported by a doctoral studentship funded by the UK Research and Innovation (UKRI).

Introduction

Questionable research practices (QRPs) can undermine the validity, reliability, and reproducibility of scientific findings, potentially compromising the integrity of the evidence base and, in turn, the quality of decisions and interventions based upon it. QRPs are without straightforward definition, but exist in an area of ethical ambiguity between outright misconduct (i.e., falsification, fabrication and plagiarism) and sound scientific practice. QRPs involve the exploitation of the inherent flexibility in the research process and may be perceived as warranted or defensible in certain contexts (John et al., 2012; Sacco et al., 2018; Fraser et al., 2018). QRPs can occur at any stage of the research process, and can be engaged in without intent to deceive (Artino et al., 2019). Broad in definition and scope, these practices include behaviours such as reporting an unexpected finding as if it were an a priori hypothesis (termed 'HARKing'; Kerr, 1998), rounding p-values to meet desired statistical thresholds ('P-hacking'), using leading questions in interviews, and excluding junior authors from a publication when their contributions have warranted authorship.

More prevalent than scientific misconduct (Xie et al., 2021), QRPs are widely discussed as possible contributors to issues of scientific replicability and reproducibility (John, Loewenstein and Prelec, 2012; Simmons et al., 2011; Wicherts et al., 2016).[8] Estimates of prevalence vary considerably across studies, though most self-report surveys and meta-analyses place prevalence levels between 12.5% and 34% (see, Martinson et al., 2005; Fanelli, 2009; Fiedler & Schwarz, 2016). However, one outlier study, based on a sample of U.S. psychologists, reported rates as high as 100% (Xie et al., 2021). Despite the variability in estimates, the fact that QRP prevalence is non-zero is a matter of significant concern, given

[8] However, it is important to note that the extent to which QRPs impact replication rates (Ulrich & Miller, 2020; Rubin, 2023; Rubin, 2025; Francis, 2012), and the degree to which these replication rates are an issue, has been subject to debate (Rubin, 2023; Rubin, 2025; Fanelli, 2018; Oberauer & Lewandowsky, 2019). That said, the author recognises that QRPs are not an issue solely because of their impact on replication failure, but because they are practices that can undermine the integrity and reliability of research more broadly.

their downstream consequences. QRPs contribute to the canonisation of erroneous scientific findings, the dissemination of inaccurate media representations, diminished public trust in science, and ultimately, the failure of policies and interventions (Jussim et al., 2019). Given their substantial impact, it is integral to identify and capture the factors and mechanisms associated with self-reported engagement in these behaviours.

The preeminent theory is that QRPs result from a misalignment between incentive structures and the ideals of good scientific practice. Academic institutions often use simple metrics, such as frequency of publication, journal impact factors and citation count when assessing for hiring, permanency, tenure and promotion (Moher et al., 2018). These systemic factors that fuel hyper-competition and perceived pressure, when combined with a publication bias that favours the aesthetic quality of results (i.e., novelty and statistical significance), motivate researchers to cut corners, cheat and ‘game the system’ to maximise their chances of publishing in high-impact journals (Martinson et al., 2005 Bouter, 2015; Edwards & Roy, 2017). Though this ‘publish or perish’ argument may serve as an overly convenient explanation that shifts attention away from individual responsibility (Yarkoni, 2018), there is some empirical evidence supporting its influence. Much of this evidence is derived from correlational research that finds self-reported publication pressure to be positively associated with self-reported self-reported QRP engagement (Brooker & Allum, 2024; Schneider et al., 2024; Gopalkrishna et al., 2022; Maggio et al., 2019). Similarly, correlational research has shown that early-career researchers, who are feasibly under greater pressure to publish to secure tenure and permanency, are more likely to engage in QRPs (Gopalakrishna et al., 2022; Brooker & Allum, 2024; Schneider et al., 2024). However, the observed effect sizes remain small.

The influence of social and psychological factors on QRP engagement remains largely unexplored. The limited research on individual differences suggests that these factors are at least as important as system-level influences in explaining QRP engagement (Brooker &

Allum, 2024; Schneider et al., 2024; Gopalakrishna et al., 2022. For example, correlational research suggests that the personality traits of agreeableness and conscientiousness are negatively associated with self-reported engagement in QRPs (Schneider et al., 2024; Giluk & Postlethwaite, 2015). Moreover, a researchers' commitment to the normative ideals of science theorised by Robert Merton (2010), have been found to be negatively associated with self-reported QRP engagement (Brooker & Allum, 2024; Gopalakrishna et al., 2022). Merton proposed that scientists adhere to a set of informal cultural norms that guide behaviour, including universalism (i.e., evaluating claims based on objective criteria rather than personal attributes), disinterestedness (pursuing research for the advancement of knowledge rather than personal gain), organised scepticism (subjecting evidence to critical scrutiny), and communalism (promoting the collective ownership and dissemination of scientific knowledge). Adherence to these normative ideals has found to be widespread, but not universal (Anderson et al., 2010; Bray & von Storch, 2017; Macfarlane & Cheng, 2008; Kim & Kim, 2018). The limited research examining the role of these scientific norms in relation to self-reported QRP engagement suggests that commitment to these ideals is among the strongest correlates of self-reported QRP engagement (Brooker & Allum, 2024; Gopalakrishna et al., 2022). However, in the studies conducted to date, scientific norms have been treated as a single, unified construct. As a result, the potentially distinct influence of each of the Mertonian norms remains largely unexplored in the existing literature.

Another potentially important yet underexamined individual-level factor is scientific orientation – namely, the extent to which researchers are motivated to falsify versus confirm hypotheses, findings and theories. A falsification-oriented approach, rooted in Popperian ideals of critical testing, may discourage selective reporting and p-hacking, while a confirmatory orientation may increase susceptibility to QRPs by motivating behaviours and practices that serve to produce publishable results. Certainly, there has been considerable debate in the sociology of science as to whether Popperian falsification truly characterises 'everyday science' (Kuhn, 1970; Barnes, 1982; Fuller, 2003). This relationship has received limited empirical attention, yet any meaningful discussion of the influence of epistemic norms on research practice must also account for how scientists approach the research process – particularly whether they are oriented toward confirming or critically testing their own and other's hypotheses, findings and theories.

Relatedly, a researcher's political orientation may feasibly influence their research practice. Research, especially in higher education settings, is primarily undertaken by individuals who are left-leaning, liberal and progressive (Gross & Fosse, 2012; Gross, 2013; Gross & Simmons, 2014; Klein, Stern, & Western, 2005; van de Werfhorst, 2019). It has been argued that this homogeneity of political orientation undermines the scientific objectivity and validity of research (Redding, 2001; Duart et al., 2015; Chambers & Schlenker, 2015), with ideologically driven errors permeating significant quantities of social research (al Gharbi, 2018). Political identities have been shown to give rise to motivated cognition (Ditto et al., 2018), which can influence decision-making at multiple stages of the research process (Duarte et al., 2015). For instance, researchers may consciously or unconsciously select analytical strategies that align with their ideological commitments, potentially increasing the likelihood of p-hacking. While some discipline-specific research has examined whether political orientation affects the replicability and robustness of scientific findings (Reinero et al., 2020), there is an absence of empirical enquiry into its relationship with self-reported engagement in QRPs.

Jussim et al. (2019) proposed that both personal motivations (e.g., career advancement, the pursuit of fame) and proximal science-specific motivations (e.g., increasing the frequency or visibility of publications, maximising impact) may influence a researcher's propensity to engage in QRPs. Researchers motivated by personal gain – such as wealth, recognition, or prestige – may be more vulnerable to compromising scientific integrity in pursuit of these ends. Likewise, strong professional motives may lead to motivated reasoning to arrive at a particular conclusion, sometimes at a cost to accuracy. In their comprehensive theoretical framework, Jussim and colleagues (2019) suggest that personal motivations can shape and reinforce science-specific motivations, thereby influencing behaviours within the research process. For instance, a personal desire for career advancement (e.g., increased salary, promotions) likely strengthens a researcher's motivation to achieve scientific success as defined by frequency of publication and citation counts, which in turn may heighten the likelihood of QRP engagement. Moreover, Jussim and colleagues speculate that the system-level causes – such as unequal incentive structures – may moderate the relationship between personal motivations and proximal science-specific motivations. While theoretically compelling, the extent to which these mechanisms influence self-reported engagement in QRPs has not been tested empirically.

The Current Study

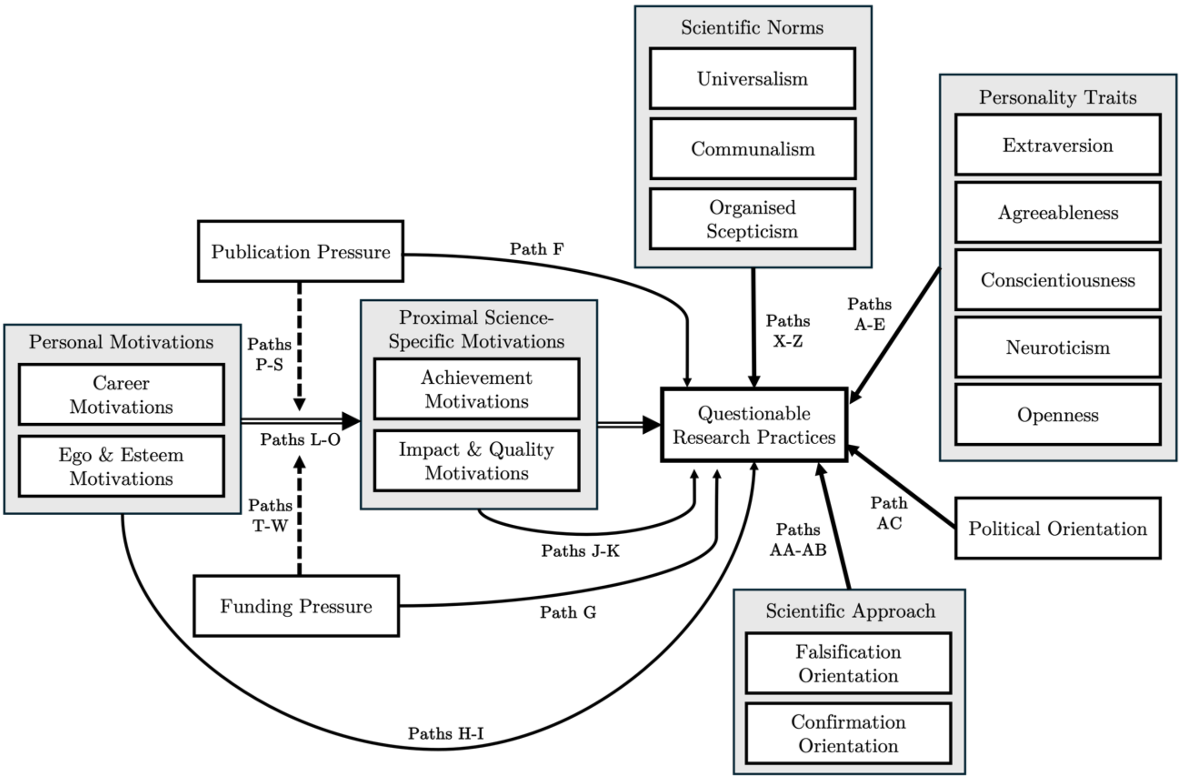
In this study, we use data from an international survey of researchers operating across disciplinary fields and countries to empirically test a conceptual model of the mechanisms and factors associated with self-reported QRP engagement. This study seeks to build on existing fragmentary evidence by synthesising the correlates of self-reported QRP engagement into a coherent and integrative empirical model. We partially build on the theoretical model developed by Jussim et al. (2019), to explore the extent to which personal motivations and proximal science-specific motivations influence self-reported QRP engagement. We hypothesise that the more individuals are motivated by career achievement (H1) and ego and esteem (H2), the greater their self-reported QRP engagement. We hypothesise that scientific achievement (e.g., obtaining a higher citation indices, publishing more frequently) is positively associated with self-reported QRP engagement (H3), while a motivation to increase the impact and quality of one's research is negatively associated with self-reported QRP engagement (H4). As proposed by Jussim et al.'s (2019), personal motivations plausibly correlate with self-reported engagement in QRPs indirectly, through their impact on more proximal, science-specific motivations. On this basis, we hypothesise that proximal science-specific motivations mediate the relationship between personal motivations and self-reported QRP engagement (H5-H8).

We hypothesise that self-reported publication pressure (H9) and funding pressure (H10) are positively associated with self-reported QRP engagement. Although these relationships feature prominently in theoretical accounts, with the influence of publication pressure receiving some empirical support, the evidence remains limited. Our study examines these associations within a broader conceptual model to assess their consistency and explanatory value. Furthermore, drawing on Jussim et al.'s (2019) theoretical framework, we examine whether publication pressure (H11-H14) and funding pressure (H15-H18) moderate the relationship between personal motivations and proximal science-specific motivations.

Building on Schneider et al. (2024) findings that conscientiousness and agreeableness are negatively correlated with self-reported QRP engagement, we hypothesise that these personality traits are negatively associated with self-reported QRP engagement (H19a and H19b). We include all five of the big five personality traits, including neuroticism, openness and extraversion, in our model for completeness. Turning to other individual level factors, we explore whether commitment to the Mertonian norms of universalism, disinterestedness and organised scepticism differentially impact self-reported QRP engagement. Moreover, we include several additional exploratory elements. We explore whether a researcher's epistemic approach – specifically, whether they primarily seek to confirm findings/theories or adopt a Popperian falsification strategy – may also influence the extent of their self-reported engagement in QRPs. We also include a measure of political orientation (left-right ideological leaning) as an exploratory correlates of self-reported engagement in QRPs. Finally, as an additional exploratory element, we include several interaction terms to identify whether publication and funding pressure moderate the influence of scientific norms and personality traits on self-reported QRP engagement.

Figure 4.1. presents our proposed theoretical model. Paths H-I correspond to the hypothesised relationships specified in H1 and H2, while paths J-K correspond to hypotheses H3 and H4. The compound line comprising paths L through O denotes the proposed mediating role of proximal science-specific motivations in the relationship between personal motivations and self-reported QRP engagement, testing in H5 through H8. Paths F and G represent the direct association of publication pressure and funding pressure on self-reported QRP engagement, testing H9 and H10, respectively. The moderating role of publication and funding pressure on the relationship between the personal motivation and proximal science-specific motivations are represented with the dashed lines (Paths P-W). This shows the hypothesised relationships corresponding to H11 through H18. Paths A–E represent our hypothesised relationships between conscientiousness (H19a), agreeableness (H19b), and self-reported engagement in QRPs, with extraversion, neuroticism, and openness included as exploratory predictors. Paths X–Z reflect our exploratory analysis of how universalism, communalism, and organised scepticism relate to self-reported engagement in QRPs. Paths AA and AB examine the influence of epistemic orientation – namely, falsification and confirmatory orientation – on self-reported QRP engagement, while Path AC explores the potential role of political orientation in shaping self-reported QRP engagement.

Figure 4.1. Conceptual model illustrating potential social-psychological correlates of self-reported engagement in questionable research practices



Note. The figure is intended to serve as a visual representation of our hypotheses and should not be construed as an overview of our analytical strategy. The compound (double-lined) arrow represents four mediation paths. The dashed-lines represent proposed moderating factors affecting the relationship between personal motivations and proximal science-specific motivations.

Data and Methods

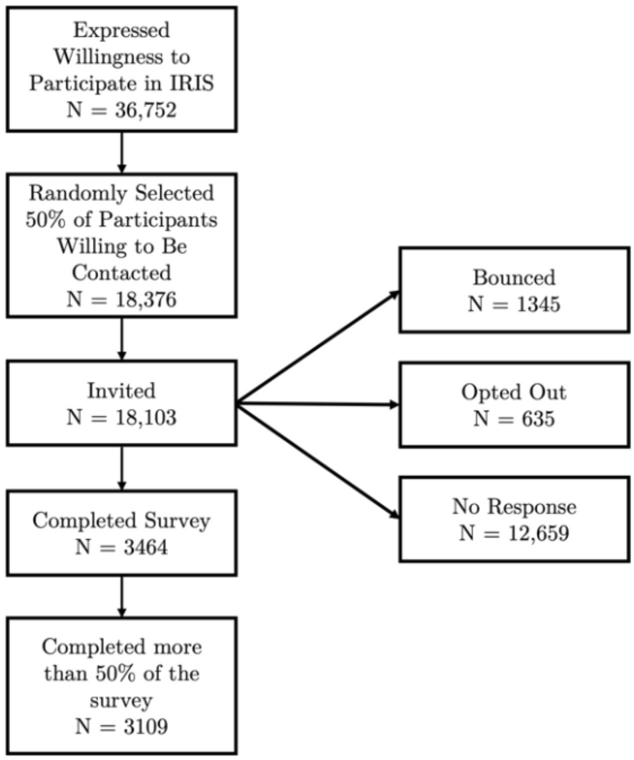
Ethical Approval

The Ethics Committee in the Department of Sociology and Criminology at the University of Essex approved the International Survey on Research Culture and Integrity (approval number: ETH2324-0515). The anonymity of the identity of respondents was ensured in accordance with UK General Data Protection Regulation (GDPR) regulation. Following merging of the International Survey on Research Culture and Integrity with the International Research Integrity Survey (IRIS; Allum et al., 2023), email addresses of respondents were removed. URLs and IP Addresses were additionally removed.

Sample

The sample was derived from a list of respondents who indicated their willingness to participate in future research after completing IRIS. Of those who indicated their willingness to participate in future studies, we randomly sampled 50% (N = 18,376). On this basis, the sample is a self-selected, non-probability volunteer sample. The original survey, from which this sample is derived, was based on a systemic, stratified probability sample of authors of research articles that were published between 2016 and 2020 and listed on Clarivate's Web of Science database. The survey protocol was pre-registered with the Centre for Open Science in November 2024 (available at: https://shorturl.at/mf9qY). We excluded individuals who had completed less than 50% of the survey (N = 355). The overall survey response rate was 19.3%, calculated using the American Association for Public Opinion Research's standard definition (AAPOR, 2023), resulting in 3109 valid survey responses. A breakdown of response statistics are shown in Figure 4.2.

Figure 4.2. Flow chart of the International Survey on Research Culture and Integrity



Study Design

The survey was a cross-sectional survey on research integrity, academic culture and academic freedom. Survey questions were meticulously developed between April and November 2024 and the survey was deployed online in November 2024 using Qualtrics – the web-based platform for survey design and administration. All researchers had previously indicated that they hold at least a master's level degree and produce research for commercial or academic organisations within the European Union, European Free Trade Area, United Kingdom or United States. No additional exclusionary criteria were applied, therefore having already met criteria for inclusion in IRIS, all individuals who received an invitation to the survey were eligible to participate. The survey was conducted entirely in English and was initially piloted with twenty-five researchers in October and November 2024. Following piloting, the survey was refined by adjusting several questions and removing items to enhance clarity, focus and reduce completion time. Between November 19th and 20th 2024, a prenotification email was sent to all individuals in the sample to inform them about the upcoming survey. Invitations to participate were sent between November 21st and 22nd. The invitation emails included information on the project, a link to the survey, an opt-out link removing recipients from further communication, and information on identity protection measures. Upon clicking the survey link, participants were greeted with the participation information sheet, followed by an informed consent page. To encourage completion, two email reminders were sent to participants who had not finished the survey, on either November 28th or 29th, and on December 6th 2024. No additional reminders were issued thereafter. The survey was officially closed to responses on December 31st, 2024. The final dataset was merged with IRIS, prior to being anonymised.

Survey Instrument

The survey included sections on demographic measures, scientific values and motivations, researcher behaviour, personality traits and academic freedom. No survey questions forced a response, and participants had the option to skip any item they were not comfortable answering.

Variables

Questionable Research Practices. We developed a repository of twenty-four questionable research practices (QRPs), drawing on previous research examining the prevalence and correlates of QRPs. Some items were adapted non-verbatim from established studies to ensure clarity and relevance across disciplinary and national contexts, while others were created specifically for this study to capture practices not addressed in prior work. Of these, eight practices pertained to 'qualitative' research, eight pertained to 'quantitative' research and the remaining eight were relevant to all research methodologies. Respondents received a maximum of twelve QRP items and a minimum of eight, with the composition of items dictated by an earlier question that asked respondents to indicate whether they primarily conduct quantitative, qualitative or mixed-methods research. If respondents indicated that they primarily adopt a quantitative or qualitative research methodology, they received the full eight-items from the respective methodological pool of items and a random subset of four items from the 'general' QRP item pool. Those who indicated that they primarily adopt a mixed methods approach received a random subset of four items from each of the methodological pools and four random items from the 'general' pool. Respondents who indicated they *do not* undertake empirical research or failed to indicate their primary methodological approach received the eight 'general' QRP items only. Descriptions of these questionable research practices can be found in Table 4.1.

Respondents were asked to indicate the frequency of which they have carried out the practice over the past three years on a five-point ordinal scale with response alternatives including never, rarely, sometimes, often, always and 'does not apply in my case'. The three-year timeframe was used to reduce recall bias and ensure prevalence could be stated in more precise terms (Xie et al., 2021; Fanelli, 2009). Measuring QRP engagement this way avoids the production of inflated and misleading prevalence estimates that measure ‘lifetime binary exposure’ (i.e., asking whether the respondent has *ever* engaged in the practice). That is, traditional surveys conflate the proportion of individual scientists who have made a single error with the true systemic frequency of the behaviour across all scientific projects. Utilising a three-year time bound that evaluates the *extent* of self-reported engagement, rather than a lifetime binary metric. Further, the data and models presented herein should be interpreted strictly as evidence of associations and behavioural patterns within the responding sample.

We removed respondents list-wise ($N = 21$) who answered fewer than four items, before deriving our dependent variables. This reduced the analytical sample to 3088 respondents. The first dependent variable is mean average self-reported QRP engagement, with mean score calculated based on all QRPs deemed to apply as indicated by the respondent. The resulting scale was standardised to z-scores, ensuring each score is mean-centered and scaled by the standard deviation. Higher scores indicate greater self-reported average engagement. This operationalisation is consistent with past research that has modelled the relationship between QRPs and various explanatory factors (e.g., Brooker & Allum, 2024; Gopalakrishna et al., 2022).

The second dependent variable is a dichotomous variable indicating whether respondents have frequently engaged in at least one QRP. Respondents indicating that they 'always' or 'often' engage in at least one of the QRP items were coded as having engaged 'frequently'. Again, operationalising the variable this way is consistent with previous research (e.g., Gopalakrishna et al., 2022).

Publication and Funding Pressure. We derived explanatory factor scales for 'publication pressure' using seven items and 'funding pressure' using five items. The question wording, response scales and mean scores stratified by field, career stage and sex, are included in Table 4.S1. Six items measuring publication pressure, and four of the items measuring funding pressure, were taken verbatim from Gopalkrishna et al. (2022). One of the items measuring funding pressure was derived and adapted from Gopalkrishna et al. (2022), while the final item measuring publication pressure was newly created for this study (see Table 4.S1). Relevant items were reverse coded, and a mean score was computed for each respondent. Higher scores indicated greater publication and funding pressure. The scales performed reasonably well in psychometric properties (with Cronbach's alpha of 0.70 for publication pressure and 0.71 for funding pressure; see Table 4.5). These explanatory scales were standardised to z-scores, ensuring each score is mean-centered and scaled by the standard deviation.

Scientific Norms. The survey included twenty items designed to measure adherence to Mertonian scientific norms. Each of the four norms – universalism, disinterestedness, organised scepticism, and communalism – were measured using five items. These items were developed based on established literature measuring these constructs (see Anderson and Louis, 1994; Anderson et al., 2010), refined through expert feedback and tested through piloting with a small sample of participants. Respondents were asked to indicate the extent to which twenty items represent the behaviours and views of researchers in the respondent's field. Respondents indicated an answer on a five-point ordinal scale, with response alternatives 'never', 'rarely', 'sometimes', 'often' and 'always'. We adopted a normative (or prescriptive) measurement strategy, where respondents were asked to evaluate how researchers should ideally behave rather than describing their actual behaviour. This approach captures researchers' perceptions of the ideal standards of scientific conduct and aligns with past scholarship that has operationalised scientific norms in a similar manner (e.g., Anderson et al., 2010).

An exploratory factor analysis (EFA) was conducted to examine the underlying structure of these normative constructs. Question wording and the complete pattern matrix for all items are included in Table 4.S2. Prior to analysis, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.76, indicating strong factorability. Bartlett's test of sphericity was significant ($X^2 = 6075.19(190)$, $p < 0.05$), confirming sufficient correlations among items for factor analysis. Given that factors may be correlated, principal axis factoring was used for factor extraction, and oblimin rotation was used to enhance interpretability.

A three-factor solution was retained based on the factor loadings and theoretical interpretability. The pattern matrix, along with factor loadings, eigenvalues, and variance explained, is presented in Table 4.2. The three retained factors explained 29.1% of the total variance, with universalism explaining 11.8%, communalism explaining 9.1% and organised scepticism explaining 8.2%. The expected factors for universalism, communalism and organised scepticism emerged, but no distinct factor corresponding to 'disinterestedness' was identified. Some items theorised to reflect disinterestedness loaded inconsistently across multiple factors, even when alternative solutions (e.g., four- and five-factor models) were explored. This suggests that disinterestedness may no longer function as a cohesive scientific norm, reflecting shifts in contemporary research practice.

Items loading onto each respective factor were averaged to create composite scores, computing a mean score for each respondent for communalism, organised scepticism, and universalism. These variables were standardised to z-scores, ensuring each score is mean-centered and scaled by the standard deviation. Higher scores indicated greater adherence to the scientific norm. The internal consistency for each factor was lower than conventional thresholds and the Cronbach's alpha threshold of 0.7 included in our pre-registration ($\alpha = 0.48 - 0.59$; see Table 4.5), but the factors were retained based on theoretical importance and the distinct clustering present in the factor structure (that is, items clearly load onto distinct factors). The lack of internal consistency likely suggests that adherence to these norms is not uniform and unidimensional, potentially reflecting shifts in contemporary research practices rather than measurement error. Moreover, the small Cronbach's alpha may result from the small number of items, as Cronbach's alpha is sensitive to scale length.

Personal, Proximal-Scientific Motivations and Scientific Approach. The survey included eighteen items measuring researcher motivations and researcher orientations. There were three broad categories of items, developed through an iterative process, incorporating insights from discussions with researchers, a review of relevant scholarship and reflections on contemporary research practices. The approach ensured that the items captured both theoretical constructs and practical considerations relevant to research motivations and epistemic conditions. These categories included personal motivations, proximal science-specific motivations and approaches to research (i.e., falsification vs. confirmation). Respondents were asked to indicate the extent to which they found each statement (e.g., 'advancing my career or achieving promotion, compared to other personal and professional motivations') motivating on a seven-point ordinal scale ranging from 0 ('Not motivating at all') to 6 ('Extremely motivating'). Full question wording and mean scores stratified by disciplinary field, career stage and sex are included in Table 4.S3.

An exploratory factor analysis was conducted to identify distinct latent constructs pertaining to personal motivations, proximal science-specific motivations and research approaches. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.80, suggesting strong factorability. Bartlett's test of sphericity was significant ($X^2 = 21032.87(153)$, $p < 0.05$) indicating that items are sufficiently correlated to justify factor analysis. A six-factor solution was extracted based on eigenvalues, factor loadings and theoretical justification. Factor loadings, eigenvalues and variance explained are presented in Table 4.3. The full pattern matrix is shown in Table 4.S4.

Table 4.1. Questionable research practice items from the international survey on research culture and integrity

	Questionable Research Practice (QRP) Question Wording
General QRPs	
QRP 1	Choosing references for citation in a way that primarily supports my own findings or beliefs.
QRP 2	Including individuals as authors on a paper who may not have contributed significantly to the research.
QRP 3	Citing literature without having read it.
QRP 4	Making adjustments to the analysis, inferences or conclusions in a manuscript that you do not believe are valid in order to satisfy reviewers.
QRP 5	Omitting citations of publications that present differing views or findings from one's own work.
QRP 6	Using content from papers under review for personal research or publication.
QRP 7	Not disclosing relevant conflicts of interest.
QRP 8	Excluding junior researchers from the list of authors on a paper despite them having made significant contributions.
Quantitative QRPs	
QRP 9	Neglecting to conduct a power analysis to determine the necessary sample size based on the intended significance level, statistical power, and anticipated effect size.
QRP 10	Not fully disclosing and making accessible all underlying data, computer codes, or syntaxes used in research on open science platforms.
QRP 11	Reporting a set of results as the complete set of analyses when other relevant analyses were also conducted.
QRP 12	Presenting an unexpected finding or exploratory analysis in a research report as if it were an a priori hypothesis.
QRP 13	Deciding not to report covariates that failed to reach statistical significance (e.g., $p < 0.05$) or some other desired statistical threshold.
QRP 14	Deciding to exclude data points after first checking the impact on statistical significance (e.g., $p < 0.05$) or some other desired statistical threshold.
QRP 15	Rounding off p-values so they appear to be below or equal to the desired alpha (e.g., reporting $p = 0.054$ as $p = 0.05$).
QRP 16	Extending data collection or manipulating data selection during an analysis with the specific intent of obtaining statistically significant results where none were previously found.
Qualitative QRPS	
QRP 17	Not maintaining comprehensive and detailed documentation of all decisions, modifications, and steps undertaken throughout the research process.
QRP 18	Not engaging in reflexivity, which involves failing to acknowledge or address my biases, perspectives, and potential influence on the research process.

QRP 19	Failing to sufficiently describe the methods of data collection and analysis, making it difficult for others to understand how conclusions were reached.
QRP 20	Phrasing questions in interviews or focus groups in a way that might influence participants towards certain responses.
QRP 21	Interpreting qualitative data in a way that confirms the analysts pre-existing ideas and beliefs, ignoring data that might suggest alternative interpretations.
QRP 22	Deliberately omitting or failing to discuss data that contradicts or does not align with the emerging themes or patterns identified in the research.
QRP 23	Presenting interview or focus group findings in a manner that may not fully reflect the original data.
QRP 24	Providing false or misleading information to participants about the purpose or nature of the study.

Table 4.2. Factor loadings for normative ideals of science

Item	Factor 1 Universalism	Factor 2 Communalism	Factor 3 Organised Scepticism	Communalities
The personal background and characteristics of a researcher should be considered when assessing the validity of their scientific claims. (Reversed)	**0.666**	-	-	0.444
The country where a researcher is based should influence a journal editor's decision to publish the researcher's work in their journal. (Reversed)	**0.451**	-	-	0.234
Researchers should take into account the authors' ideological positions when evaluating the validity of the findings or conclusions presented in a scientific article. (Reversed)	**0.605**	-	-	0.354
The personal beliefs and values of researchers should play a role in the interpretation of data or outcomes in research. (Reversed)	**0.361**	-	-	0.171
Researchers should aim to consider all new evidence, hypotheses, and theories, even those that challenge or contradict their own work.	-	-	**0.679**	0.457
Researchers should reject scientific findings solely because they contradict well-established knowledge. (Reversed)	-	-	**0.368**	0.152
Researchers should thoroughly scrutinise the methodology and inferences of a published study before making judgments about the validity and reliability of its findings.	-	-	**0.483**	0.238
Research findings should be a communal resource and not proprietary knowledge owned by individuals or institutions.	-	**0.406**	-	0.192
Researchers should be obligated to upload data to public databases and repositories.	-	**0.715**	-	0.503
Researchers should withhold raw data and methodologies to prevent others from duplicating or challenging their work. (Reversed)	-	**0.372**	-	0.241
Eigenvalue	1.176	0.908	0.815	
% Variance	0.118	0.091	0.082	
Cronbach's Alpha	0.59	0.52	0.48	

Factor loadings below 0.30 are suppressed for clarity.
EFA was conducted using maximum likelihood with oblique (oblimin) rotation.
Communalities (h^2) and eigenvalues for each factor are provided to indicate the proportion of variance explained.
For the complete pattern matrix, including suppressed loadings, see supplementary material.

The two personal motivation factors comprised a 'career growth' motivation, reflecting a drive for career growth and security, and an 'ego & esteem' motivation, reflecting a desire to obtain fame, avoid external perceptions of failure and achieve political or social goals. The two proximal science-specific motivations comprised a motivation for 'impact and quality', reflecting a drive for scientific and societal impact, and a drive for 'academic/scientific achievement', reflecting a motivation to improve one's scientific success (e.g., number of publications, citation index). Impact & quality motivation and ego & esteem motivation exhibited lower reliability, with a Cronbach's α of 0.52 and 0.58 respectively. This likely indicates some heterogeneity in these motivations or overlapping dimensions not fully captured by the factor structure. Career growth motivation and achievement motivation demonstrated good internal consistency ($\alpha = 0.74$ and 0.83, respectively), supporting their validity as distinct and reliable constructs. A clear 'falsification' vs. 'confirmation' distinction emerged, with both orientations showing strong internal consistency ($\alpha = 0.83$ and 0.81, respectively). Falsification orientation reflects an emphasis on critically testing and potentially refuting theories and findings, including one's own. Confirmation orientation reflects a preference for reinforcing one's previous research and theories.

The retained factors accounted for 51% of total variance, with individual factor contributions ranging from 5% to 11.6%. Items loading onto each factor were averaged to create composite scores, computing a mean score for each respondent for each of the motivations and orientations. Higher scores indicated greater motivation or stronger orientation. These scales were standardised to z-scores, ensuring each score is mean-centered and scaled by the standard deviation.

Personality Traits. We used the Ten-Item Personality Inventory (TIPI) (Gosling et al., 2003) to measure the Big Five personality traits, including extraversion, agreeableness, conscientiousness, emotional stability (opposite of neuroticism), and openness to experience.

The TIPI was chosen for its brevity and suitability for large-scale surveys while effectively capturing the core dimensions of the Big Five personality traits. Despite its short length, the inventory has demonstrated acceptable convergent validity with established Big Five measures (Gosling, et al., 2003).

The measure consists of ten items, with two per trait (one positively and one negatively worded), such as *"extraverted, enthusiastic"* (for extraversion). Respondents were asked to indicate the extent to which they agreed or disagreed with each descriptor pair on a 7-point Likert scale (1 = Disagree strongly, 7 = Agree strongly). Negatively worded items were reverse-coded, and scores for each trait were computed by averaging responses for the two relevant items. Finally, these scores were standardised to z-scores, ensuring that each variable was mean-centered with a standard deviation of one.

Demographic Variables. In this survey, we did not re-collect extensive demographic data, as these variables were obtained in IRIS (2021). Given that our sample is derived from those who completed IRIS, responses were matched to previously collected demographic data. On this basis, only an updated measure of career stage, and a new measure of methodological approach were collected. Career stage was measured by asking respondents to select from early career, mid-career, late career or retired. This variable was recoded into three dummy variables, with 'mid-career' as the reference category, chosen because it represented the largest category. Respondents were initially asked whether they engage in empirical research. If they indicated that they engage in empirical research, they were asked whether they primarily use a quantitative, qualitative or a mixed-method approach. This variable was recoded into two dummy variables, with quantitative as the reference category. Quantitative researchers were used as the reference category, as many QRPs (e.g., p-hacking, HARKing) are most closely associated with quantitative research traditions, making this a substantively meaningful baseline for comparison. We used demographic measures of sex, discipline and country of employment, as measured in IRIS (2021).

In IRIS, each respondent reported their sex, with response categories male, female and prefer not to say. Those selecting 'prefer not to say' (*N* = 38) were removed list-wise from the analysis, reducing our analytical sample down to 3050. This variable was recoded into a dummy variable, with males as the reference category, as they represented the largest group in the sample. Disciplinary field was measured by asking respondents to report their main disciplinary field. There responses were assigned to four broad categories, including social sciences, humanities, medical sciences (including biomedical sciences) and the natural sciences (including technical sciences). This variable was recoded into three dummy variables, with natural sciences as the reference category, as they represent the largest group in the sample. Respondents were additionally grouped into 'United States of America', 'Europe' and 'other', based on their self-reported country of employment. This variable was recoded into a dummy variable with 'Europe' as the reference category. We additionally asked respondents to indicate where they position themselves on a left-right political scale, ranging between 0-6 (7-point Likert scale), with the label 'far left' anchored at 0, 'far right' anchored at 6 and 'centre' anchored at 3. The left-right political position variable was standardised to z-scores, with mean zero and standard deviation one.

Survey Weights. Survey weights were constructed to adjust for differential selection and nonresponse bias. Respondents who indicated willingness to participate were randomly selected with a fixed probability of 0.5, and the likelihood of providing a usable response (defined as completing at least 50% of the survey) was estimated using a logistic regression model that included country, field group, field, and their interactions. Raw weights were derived as the inverse of the product of this fixed selection probability and the predicted response probability, and these weights were subsequently normalised to have a mean of one.

This approach ensures that our subsequent analyses accurately reflect the intended survey design while mitigating biases associated with nonresponse. The code used to produce survey weights is included in the R code uploaded to OSF.

Missing Data

Data was found to be Missing at Random (MAR), therefore Multiple Imputation with mice in R (version 2023.09.0+463) was used to produce fifteen complete datasets (van Buuren & Groothius-Oudshoorn, 2011). Predictive mean matching was used for continuous variables, proportional odds logistic regression for ordinal variables (including Likert scales) and multinomial logistic regression was used for unordered categorical variables. Models were fit to the fifteen imputed datasets, and the results were combined into a single inference. To account for uncertainty all estimates were combined using Rubin's rules (Rubin, 1987). Complete case analyses were conducted as a sensitivity check for all models.

Analysis Strategy

Our modelling strategy is as follows. In Models 1A and 2A, we specify a linear regression model and a logistic regression model, regressing our primary explanatory factors of interest – publication pressure (Path F), funding pressure (Path G), personal motivations (Paths H-I), proximal science-specific motivations (Paths 3-4), epistemic orientation (Paths AA-AB), and personality traits (Paths A-E) – on Mean QRP and Any Frequent QRP. This allows us to test H1 to H10, while additionally conducting our pre-registered exploratory analysis investigating whether adherence to communalism, universalism and organised scepticism correlate with self-reported engagement in QRPs (Paths X-Z). We include measures of each of the Big Five personality traits on self-reported QRP engagement. We do not include a measure of disinterestedness, contrary to our pre-registration, as our factor analysis did not identify it as a unidimensional construct. We also include a measure of left-right political position in Models 1A to 2A as an additional exploratory analysis (Path AC).

In Models 1B and 2B, we include several interaction terms in exploratory linear and logistic regression models, allowing us to ascertain whether publication and funding pressure potentially moderate the influence of organised scepticism, universalism and communalism on self-reported QRP engagement. We additionally include interaction terms to ascertain whether publication pressure moderates the influence of conscientiousness, neuroticism, openness and agreeableness on self-reported QRP engagement.

Table 4.3. Factor Loadings for Personal Motivations, Proximal-Science Motivations, and Epistemic Approach Orientation

Item	Factor 1 Achievement Motivation (Proximal Science-Specific)	Factor 2 Falsification Orientation (Approach)	Factor 3 Career Growth Motivation (Personal)	Factor 4 Confirmation Orientation (Approach)	Factor 5 Ego & Esteem Motivation (Personal)	Factor 6 Impact & Quality Motivation (Proximal Science-Specific)	Communalities
Obtaining a higher citation index (scimotcite)	**0.684**	-	-	-	-	-	0.548
Increasing my number of publications (scimotpubn)	**0.855**	-	-	-	-	-	0.685
Increasing my likelihood of publication in a prestigious outlet (scimotpres)	**0.764**	-	-	-	-	-	0.673
Increasing my opportunities for research funding (scimotfund)	**0.434**	-	-	-	-	-	0.471
Disconfirming my past findings (scimotdisfin)	-	**0.776**	-	-	-	-	0.761
Disconfirming my own theories (scimotdisthe)	-	**1.030**	-	-	-	-	0.972
Disconfirming competing theories or findings (scimotdiscomp)	-	**0.446**	-	-	-	-	0.341
Advancing my career or achieving promotion, compared to other personal and professional motivations (motcareer)	-	-	**0.638**	-	-	-	0.563
Securing financial benefits (e.g., an increase in pay) or rewards for myself, compared to other personal and professional motivations (motfinan)	-	-	**0.842**	-	-	-	0.663
Ensuring my job security or reducing the risk of job loss, compared to other personal and professional motivations (motjobsec)	-	-	**0.566**	-	-	-	0.346
Confirming my past findings (scimotconfin)	-	-	-	**0.998**	-	-	0.995
Confirming my own theories (scimotconthe)	-	-	-	**0.535**	-	-	0.522
Achieving my social or political goals, compared to other personal and professional motivations (motsocgoal)	-	-	-	--	**0.346**	-	0.183
Obtaining respect or fame, compared to other personal and professional motivations (motresfame)	-	-	-	-	**0.999**	-	0.995
Avoiding being perceived as a failure or unsuccessful in my field, compared to other personal and professional motivations (motfail)	-	-	-	-	**0.315**	-	0.245

Producing new and exciting research findings (scimotnew)	-	-	-	-	-	**0.621**	0.413
Maximising the impact of my research outside of academia (scimotimpact)	-	-	-	-	-	**0.420**	0.204
Ensuring the reproducibility and accuracy of my research findings (scimotrepro)	-	-	-	-	-	**0.479**	0.359
Eigenvalue	2.079	1.912	1.606	1.401	1.264	0.895	
% Variance	0.115	0.106	0.089	0.078	0.070	0.050	

Factor loadings below 0.30 are suppressed for clarity. EFA was conducted using maximum likelihood with oblique (oblimin) rotation.
Communalities (h^2) and eigenvalues for each factor are provided to indicate the proportion of variance explained.
For the complete pattern matrix, including suppressed loadings, see supplementary material.

Models 3A and 3B regress publication pressure, funding pressure, and our two personal motivation variables, career growth and ego-esteem, on our two proximal-science specific motivations, academic achievement and impact and quality motivations. To test whether publication pressure and funding pressure moderate the relationship between our personal motivations and proximal-science specific motivations (Paths P-W), we include four interaction terms. Models 3A and 3B allow us to test H12-H19.

To examine whether proximal science-specific motivations mediate the relationship between personal motivations and QRPs (Paths L-O), we conduct a mediation analysis using Structural Equation Modelling (SEM). The SEM was analysed in R using the lavaan package. Models 4A, 4B and 4C include career achievement and impact and quality motivations as mediators, career growth and ego and esteem motivations as independent variables and Mean QRP as the dependent variable, allowing us to test H20-H23.

All multivariable models contain measures of our five demographic characteristics, including disciplinary field, career stage, methodological approach, sex and country of employment. To ensure robustness bootstrapped confidence intervals are produced for all coefficient estimates across all models. All models include survey weights.

Results

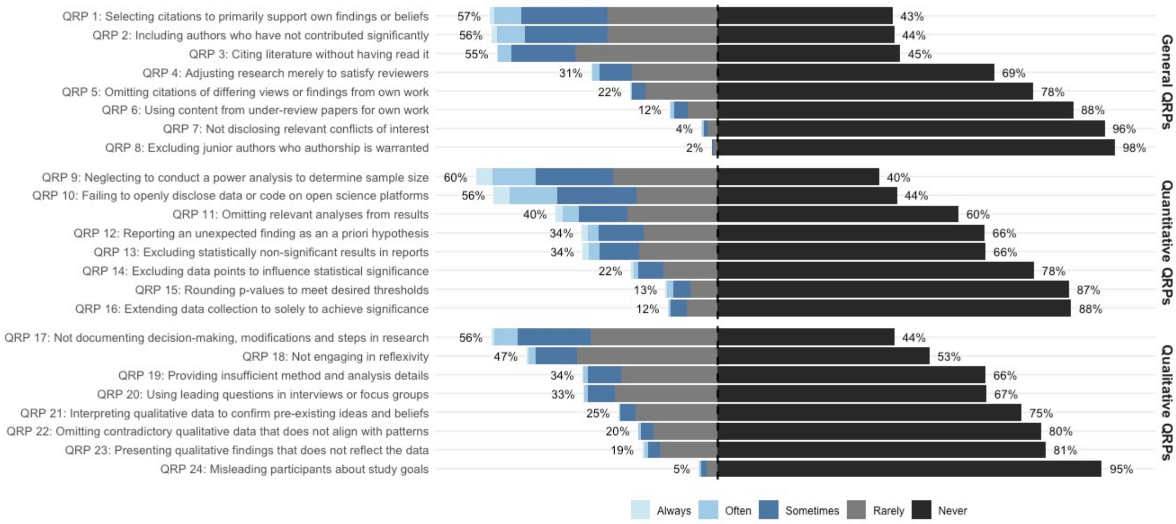
Descriptive Statistics

A small majority of our respondents were male (57%), 37.1% primarily research in the natural sciences, 14.1% in the medical sciences, 31% in the social sciences and 17.8% in the humanities and arts. 19.6% were early-career, 41.1% were mid-career, 29% were late career and 10.2% were retired. Moreover, 4.3% were based in the United States, 84% were based in Europe and 11.4% were based elsewhere. Descriptive statistics of demographic variables were calculated for each of the 15 imputed datasets and pooled by averaging across the imputed datasets.

Questionable Research Practices

Figure 4.3 shows the percentage of observed answer categories for the individual QRP items, grouped by quantitative, qualitative and general methodological approaches. A majority of researchers have indicated that at some point over the past three years, they have 'selectively cited' to support one's own findings or beliefs, cited literature without having read it and gifted authorship where authorship was not warranted. Comparably, a very small proportion of researchers report to have excluded junior authors from publications where authorship was warranted. A majority of quantitative researchers have reported that they have neglected to conduct *a priori* power analyses to determine sample size and have failed to make data, code and syntax accessible to others, contrary to optimal transparency standards. Likewise, a sizable minority of quantitative researchers reported that at some point over the past three years, they have failed to comprehensively report all relevant analyses when reporting results. Engaging in a group of practices that are collectively described as 'p-hacking' (e.g., QRPs 14-16) appear to be engaged in less frequently.

Figure 4.3. Percentage of observed answer categories of QRPs across 3088 respondents (N = 3055)



Turning to qualitative research, most qualitative researchers report to have not consistently maintained comprehensive notes on decision-making, modifications and steps in the research process, while a large minority report that they have at some point over the past three years not reflexively engaged with their research. Misleading participants about the true nature of research is less common. Percentages for each answer category and confidence intervals for all QRP items are shown in Table 4.S5.

Table 4.4 shows descriptive statistics for the unstandardised dependent variables, Mean QRP and Any Frequent QRP, stratified by disciplinary field, career stage and sex. Mean self-reported QRP engagement averages at the lower end of the scale, indicating lower average self-reported engagement. There is limited variability in average self-reported QRP engagement across disciplinary fields, career stages and between males and females. The proportion of researchers reporting self-reported frequent engagement in at least one QRP is similar across disciplinary fields (between 30-33%), with the exception of the humanities where only 17% of researchers report to having engaged in at least one QRP frequently. There is considerable variability in the proportion of researchers reporting engagement in at least one relevant QRP frequently across career stages, with a smaller proportion of late-career and retired researchers engaging in these practices, in comparison to early-career and mid-career researchers. There is limited variability in the proportion of researchers engaging at least one relevant QRP frequently between males and females.

Table 4.4. Mean average questionable research practice (QRP) (SD) and any frequent QRP stratified by disciplinary field, career stage and sex

	Disciplinary field				Career stage				Sex		Overall
Explanatory Factor Scale	**Natural sciences**	**Medical sciences**	**Social sciences**	**Humanities**	**Early career**	**Mid career**	**Late career**	**Retired**	**Male**	**Female**	
Mean QRP	1.49 (0.36)	1.49 (0.39)	1.5 (0.37)	1.41 (0.35)	1.57 (0.43)	1.49 (0.37)	1.41 (0.33)	1.42 (0.36)	1.49 (0.37)	1.47 (0.38)	1.48 (0.37)
Any Frequent QRP (% Having Engaged)	30.54	32.24	31.60	17.02	36.94	31.12	22.23	21.83	28.77	28.31	28.73

n = 3050; Mean QRP is unstandardised and computed based on 4-12 QRP items with not applicable responses removed.
Any frequent QRP is at least one of the 12 QRPs being often or always being engaged in.

Table 4.5 presents the descriptive statistics for all explanatory factor scales, disaggregated by disciplinary field, career stage, and sex. Across these subgroups, both publication and funding pressure average at the higher end of the scale, indicating that researchers perceive substantial systemic pressures. Similarly, average scores for commitment to the Mertonian norms of universalism, communalism, and disinterestedness are also located towards the upper end of the scale, reflecting a relatively strong endorsement of these normative ideals across fields, career stages, and sexes. Averages for each of the motivation variables are situated at the higher end of the scale, with some variability across disciplinary fields, career stages and sex. On average, researchers are more motivated to enhance the impact and quality of their research, compared to all other motivations, while researchers report a stronger orientation toward a confirmatory approach rather than toward the falsification of research.

Regression Analyses

Tables 4.6 shows the multivariate regression estimate for the five demographic characteristics. The multivariate regression results for Models 1A, 1B, 2A and 2B are shown in Table 4.7. Regression estimates for each individual self-reported QRP item are presented in the supplementary material (Table 4.S6**)**. Models 1A and 1B show coefficient estimates for mean frequency of self-reported engagement in Questionable Research Practices (QRPs), while Models 2A and 2B provide odds ratios for any frequent self-reported QRP engagement. Interaction terms are incorporated into Models 1B and 2B to assess potential moderating associations. All models contain the five background characteristics (disciplinary field, career stage, methodological approach and country).

Table 4.5. Mean scores (standard deviations) of explanatory factors and scales stratified by disciplinary field, career stage and sex

	Disciplinary field				**Career stage**				**Sex**		**Overall**
Explanatory Factor Scale	**Natural sciences**	**Medical sciences**	**Social sciences**	**Humanities**	**Earl career**	**Mid career**	**Late career**	**Retired**	**Male**	**Female**	
Publication Pressure	4.37 (0.03)	4.17 (0.05)	4.37 (0.03)	4.33 (0.04)	4.64 (0.04)	4.44 (0.03)	4.16 (0.03)	3.82 (0.05)	4.24 (0.02	4.47 (0.03)	4.33 (0.02)
Cronbach's Alpha: 0.70											
Funding Pressure	4.84 (0.04)	4.57 (0.06)	4.34 (0.04)	4.22 (0.06)	5.03 (0.05)	4.68 (0.04)	4.23 (0.04)	3.90 (0.07)	4.41 (0.03)	4.70 (0.04)	4.54 (0.02
Cronbach's Alpha: 0.71											
Universalism	4.27 (0.02)	4.03 (0.03)	3.97 (0.02)	3.91 (0.03)	4.13 (0.03)	4.12 (0.02)	4.06 (0.02)	3.87 (0.04)	4.08 (0.02)	4.09 (0.02)	4.08 (0.01
Cronbach's Alpha: 0.59											
Communalism	4.06 (0.02)	3.95 (0.03)	3.92 (0.02)	3.96 (0.03)	4.00 (0.03)	3.97 (0.02)	3.99 (0.02)	4.00 (0.04)	4.04 (0.02)	3.90 (0.02)	3.98 (0.01
Cronbach's Alpha: 0.52											
Organised Scepticism	4.67 (0.01)	4.78 (0.02)	4.71 (0.01)	4.76 (0.02)	4.68 (0.02)	4.71 (0.01)	4.74 (0.01)	4.72 (0.02)	4.69 (0.01)	4.76 (0.01)	4.72 (0.01
Cronbach's Alpha: 0.48											
Achievement Motivation	3.71 (0.04)	3.83 (0.07)	3.69 (0.04)	3.43 (0.06)	3.92 (0.05)	3.76 (0.04)	3.61 (0.05)	3.02 (0.09)	3.59 (0.03)	3.80 (0.04)	3.67 (0.02
Cronbach's Alpha: 0.83											
Impact & Quality Motivation	4.76 (0.03)	4.86 (0.05)	4.67 (0.03)	4.60 (0.05)	4.66 (0.04)	4.77 (0.03)	4.72 (0.03)	4.58 (0.06)	4.61 (0.02)	4.85 (0.03)	4.72 (0.02
Cronbach's Alpha: 0.52											
Ego & Esteem Motivation	2.80 (0.04)	2.94 (0.07)	3.05 (0.04)	3.08 (0.06)	2.5 (0.05)	2.95 (0.04)	2.96 (0.04)	2.88 (0.09)	2.85 (0.03)	3.07 (0.04)	2.95 (0.02
Cronbach's Alpha: 0.58											
Career Growth Motivation	3.41 (0.04)	3.40 (0.07)	3.52 (0.05)	3.56 (0.06)	3.93 (0.05)	3.60 (0.04)	3.31 (0.05)	2.52 (0.08)	3.34 (0.03)	3.64 (0.04)	3.47 (0.03
Cronbach's Alpha: 0.74											
Falsification Orientation	2.86 (0.05)	3.05 (0.07)	2.66 (0.05)	2.80 (0.07)	2.79 (0.06)	2.76 (0.04)	2.85 (0.05)	2.98 (0.09)	2.91 (0.04)	2.70 (0.04)	2.82 (0.03
Cronbach's Alpha: 0.83											
Confirmation Orientation	3.62 (0.05)	3.70 (0.07)	3.10 (0.05)	3.40 (0.07)	3.50 (0.07)	3.34 (0.05)	3.45 (0.05)	3.67 (0.09)	3.41 (0.04)	3.49 (0.04)	3.44 (0.03
Cronbach's Alpha: 0.81											
Extraversion	3.98 (0.04)	4.21 (0.07)	4.26 (0.05)	4.22 (0.06)	4.00 (0.06)	4.10 (0.04)	4.28 (0.05)	4.23 (0.08)	4.00 (0.04)	4.33 (0.04)	4.14 (0.03
Agreeableness	4.66 (0.04)	4.88 (0.05)	4.82 (0.04)	4.99 (0.05)	4.75 (0.05)	4.78 (0.03)	4.85 (0.04)	4.88 (0.07)	4.66 (0.03)	4.99 (0.03)	4.80 (0.02
Conscientiousness	5.68 (0.04)	5.97 (0.05)	5.81 (0.04)	5.95 (0.05)	5.63 (0.05)	5.78 (0.03)	5.93 (0.04)	5.96 (0.06)	5.71 (0.03)	5.95 (0.03)	5.81 (0.02
Emotional Stability (Neuroticism)	5.08 (0.04)	5.47 (0.06)	5.07 (0.04)	5.10 (0.06)	4.85 (0.06)	5.10 (0.04)	5.29 (0.05)	5.41 (0.07)	5.24 (0.03)	5.02 (0.04)	5.14 (0.02
Openness	5.72 (0.03)	5.75 (0.05)	5.67 (0.03)	5.92 (0.04)	5.57 (0.04)	5.68 (0.03)	5.88 (0.03)	5.97 (0.05)	5.73 (0.02)	5.77 (0.04)	5.75 (0.02

Explanatory factor scales are unstandardised.

Publication pressure and funding pressure are composite scales ranging from 1 to 5 with higher scores indicating greater pressure. Universalism, organised scepticism and communalism are composite scales ranging from 1 to 5 with higher scores indicating greater commitment. Personality traits were measured using composite scales ranging from 1 to 6, where higher scores indicate a stronger presence of the trait. Achievement motivation, impact and quality motivation, ego and esteem motivation, career growth motivation, falsification orientation and confirmation orientation are composite scales ranging from 1 to 6 with higher scores indicating greater motivation/orientation.

Table 4.6 shows that researchers in the humanities report to engage in fewer QRPs on average, and are less likely to report frequent QRP engagement, compared to their counterparts in the natural sciences. Average self-reported QRP engagement reduces as researchers move through their careers, with those earlier in their career reporting greater average self-reported QRP engagement compared to researchers in the middle of their career, while researchers who are later in their career report lower average self-reported QRP engagement compared to their mid-career counterparts. Self-reported QRP engagement averages lower among qualitative and mixed-method researchers compared to quantitative researchers, while quantitative researchers are also more likely to have engaged in a QRP frequently compared to their qualitative and mix-methods counterparts. We find no statistically significant sex or country differences.

Table 4.6. Regression Coefficients (95% confidence interval) of overall mean questionable research practices (QRP) and any frequent QRP stratified by demographic characteristics

	Overall Mean QRP Linear Regression Model Coefficient (95% CI)	Any Frequent QRP Logistic Regression Model OR (95% CI)
Disciplinary field (reference: natural sciences)		
Medical sciences	0.03 (-0.08, 0.14)	1.15 (0.88, 1.50)
Social sciences	-0.02 (-0.11, 0.07)	1.10 (0.88, 1.37)
Humanities	**-0.18 (-0.28, -0.08)**	**0.61 (0.45, 0.83)**
Career stage (reference: mid-career)		
Early career	**0.12 (0.02, 0.22)**	1.19 (0.96, 1.47)
Late career	**-0.14 (-0.22, -0.07)**	**0.68 (0.55, 0.85)**
Retired	-0.09 (-0.22, 0.03)	0.73 (0.53, 1.02)
Methodological approach (reference: quantitative)		
Qualitative	**-0.10 (-0.16, -0.04)**	**0.69 (0.59, 0.80)**
Mixed-methods	**-0.04 (-0.07, -0.01)**	**0.90 (0.84, 0.96)**
Sex (reference: male)		
Female	-0.04 (-0.12, 0.03)	0.95 (0.79, 1.14)
Country (reference: Europe)		
United States	-0.01 (-0.16, 0.14)	1.33 (0.87, 2.03)
Other	-0.05 (-0.17, 0.07)	0.98 (0.74, 1.29)

n = 3050; Estimates from Models 1B and 2B, including all independent variables and interactions. Coefficients and confidence intervals (CIs) are based on pooled estimates across 15 imputed datasets, using Rubin's rules. CIs are bootstrapped. Sampling weights were included in all models.
Overall mean QRP was computed based on 4-12 QRP items with not applicable responses removed.
Any frequent QRP is computed based on at least one of the 12 QRPs being 'often' or 'always' engaged in.
Bolded figures indicate statistical significance at the **0.05** alpha level.

Table 4.7 shows that one standard deviation increase in publication pressure is associated with a 0.11 ($p < 0.05$) standard deviation increase in average self-reported QRP engagement. Likewise, we find that for every standard deviation increase in publication pressure, the odds of engaging in at least one QRP frequently increases by 1.24 ($p < 0.05$). Contrary to H10, we find no statistically significant association between funding pressure and

a) average self-reported QRP engagement and b) any frequent self-reported QRP engagement. A greater commitment to universalism, communalism and organised scepticism is associated with less average self-reported QRP engagement. Universalism and organised scepticism were the strongest correlates of self-reported QRP engagement across all models, with a one standard deviation increase in commitment associated with a 0.14 and 0.15 ($p < 0.05$) standard deviation decrease in average self-reported QRP engagement, respectively. Greater commitment to organised scepticism was additionally associated with 0.89 ($p < 0.05$) decreased odds of engaging in at least one QRP frequently. There is a statistically significant interaction between publication pressure and universalism in predicting self-reported QRP engagement ($\beta = –0.04$, $p < 0.05$). The negative coefficient indicates that the association between publication pressure and average self-reported QRP engagement weakens as commitment to the normative ideal of universalism increases. That is, researchers who report a greater commitment to universalism are less susceptible to the influence of publication pressure on self-reported QRP engagement. However, the effect size is very small ($\beta = -0.04$).

A standard deviation increase in ego and esteem motivation is associated with a 0.11 ($p < 0.05$) standard deviation increase in average self-reported QRP engagement. Moreover, where a researcher is more motivated by impact and quality, average self-reported QRP engagement is lower, with a 0.11 ($p < 0.05$) standard deviation decrease in average self-reported QRP engagement for a one standard deviation increase on the impact and quality motivation scale. Conversely, where a researcher is more motivated by academic achievement (i.e., higher citation indices, more frequent publication) average self-reported QRP engagement is greater, with a 0.07 ($p < 0.05$) standard deviation increase in average self-reported QRP engagement for each standard deviation increase on the achievement motivation scale. A researcher primarily orientated towards 'falsifying' theories and findings is associated with a decreased average self-reported QRP engagement ($\beta = –0.06$, $p < 0.05$). Moreover, for each standard deviation increase on the falsification scale the odds of engaging in a QRP frequently decreases by a factor of 0.85 ($p < 0.05$). We find no statistically significant relationship between a confirmatory research approach and self-reported QRP engagement. Higher levels of agreeableness, conscientiousness, emotional stability (opposite to neuroticism) and openness are associated with lower average self-reported QRP engagement. Relatedly, the odds of engaging in any frequent QRP decrease by 0.90 ($p < 0.05$) for each unit increase in agreeableness. While we find no statistically significant direct relationship between extraversion and self-reported QRP engagement, the interaction between extraversion and publication pressure is significant (OR = 1.11, $p < 0.05$). This indicates that extraversion to some extent moderates the relationship between publication pressure and frequent QRP engagement. That is, the positive association of publication pressure and frequent self-reported QRP engagement is stronger among researchers who are more extraverted.

Moderation Analyses

Table 4.8 shows the results of the regression analyses for Models 3A and 3B, which examine whether publication pressure and funding pressure moderate the relationship between the personal motivations (i.e., career growth, and ego and esteem) and proximal science-specific motivations (i.e., achievement, and impact and quality). We find no statistically significant moderating role of publication pressure nor funding pressure on the relationship between the personal motivations and proximal science-specific motivations. However, we find a statistically significant association between publication pressure, funding pressure and the personal motivations on both the proximal-science specific motivations of achievement and impact/quality.

Mediation Analyses

Tables 4.9-4.11 show the results from Models 4A-4C, which are our structural equation models examining the theorised relationships among personal motivations and proximal science-specific motivations, and average self-reported QRP engagement. Specifically, Table 4.9 shows the fit statistics for each structural model, while Table 4.10 shows the direct associations and Table 4.11 shows the indirect associations. The personal motivations (PM) of career growth and ego and esteem are included as predictors, while the proximal science-specific motivations of achievement motivation and impact and quality are included as mediators (M). Average self-reported QRP engagement is included as the outcome. We allowed for the possibility of direct associations of personal motivations on QRPs (following the earlier regression analyses), but our primary interest was the mediating role of proximal science-specific motivations on the relationship between personal motivations and QRPs. The SEM models were estimated to test H12-H19. All variables were treated as manifest, and the models were estimated using maximum likelihood estimation. Model 4B removed non-significant demographic predictors, while the final model (Model 4C) added a covariance between the two proximal-science specific mediators.

Table 4.7. Regression Coefficients (95% confidence interval) for overall mean questionable research practices (QRP) and any frequent QRP stratified by explanatory variables

	Overall Mean QRP		Any Frequent QRP	
	Linear Regression Model 1A Coefficient (95% CI)	**Linear Regression Model 1B Coefficient (95% CI)**	**Logistic Regression Model 2A OR (95% CI)**	**Logistic Regression Model 2B OR (95% CI)**
Publication Pressure	**0.12 (0.08, 0.15)**	**0.11 (0.08, 0.15)**	**1.24 (1.12, 1.37)**	**1.24 (1.12, 1.38)**
Funding Pressure	-0.01 (-0.05, 0.03)	-0.01 (-0.05, 0.03)	1.05 (0.94, 1.16)	1.05 (0.94, 1.17)
Scientific Norms				
Universalism	**-0.15 (-0.19, -0.10)**	**-0.14 (-0.18, -0.1)**	**0.9 (0.82, 0.99)**	0.91 (0.83, 1)
Communalism	-0.04 (-0.08, 0.00)	**-0.03 (-0.07, 0)**	0.93 (0.85, 1.02)	0.93 (0.84, 1.02)
Organised Scepticism	**-0.15 (-0.2, -0.1)**	**-0.15 (-0.2, -0.1)**	**0.91 (0.83, 0.99)**	**0.89 (0.82, 0.98)**
Personal Motivations				
Career Growth	-0.03 (-0.07, 0.02)	-0.03 (-0.07, 0.01)	0.9 (0.81, 1.01)	0.9 (0.8, 1.01)
Ego and Esteem	**0.11 (0.07, 0.15)**	**0.11 (0.07, 0.15)**	1.09 (0.98, 1.20)	1.08 (0.98, 1.2)
Proximal Science-Specific Motivations				
Impact and Quality	**-0.1 (-0.15, -0.06)**	**-0.11 (-0.15, -0.06)**	0.92 (0.83, 1.02)	0.92 (0.83, 1.03)
Academic Achievement	**0.06 (0.02, 0.11)**	**0.07 (0.02, 0.11)**	1.12 (0.99, 1.27)	1.13 (1.00, 1.28)
Research Approach				
Falsification	**-0.06 (-0.11, -0.01)**	**-0.06 (-0.11, -0.01)**	**0.85 (0.75, 0.96)**	**0.85 (0.75, 0.96)**
Confirmation	0.04 (-0.01, 0.09)	0.03 (-0.01, 0.08)	1.03 (0.91, 1.18)	1.04 (0.91, 1.18)

Personality Traits				
Extraversion	0.03 (-0.01, 0.06)	0.02 (-0.01, 0.06)	1.03 (0.94, 1.12)	1.02 (0.92, 1.12)
Agreeableness	**-0.04 (-0.07, 0.00)**	**-0.03 (-0.07, 0.00)**	**0.9 (0.82, 0.98)**	**0.9 (0.82, 0.98)**
Conscientiousness	**-0.07 (-0.11, -0.03)**	**-0.07 (-0.11, -0.03)**	1.05 (0.96, 1.15)	1.05 (0.96, 1.16)
Neuroticism	**-0.08 (-0.11, -0.04)**	**-0.08 (-0.11, -0.04)**	**0.89 (0.82, 0.98)**	0.92 (0.83, 1.01)
Openness	**-0.08 (-0.12, -0.04)**	**-0.08 (-0.12, -0.04)**	0.95 (0.86, 1.04)	0.95 (0.86, 1.05)
Political Orientation	-0.03 (-0.07, 0.02)	-0.03 (-0.07, 0.01)	0.99 (0.91, 1.1)	1.00 (0.91, 1.10)
Universalism * Funding Pressure		0.01 (-0.02, 0.05)		0.93 (0.84, 1.03)
Universalism * Publication Pressure		**-0.04 (-0.08, 0.00)**		0.97 (0.88, 1.06)
Organised Scepticism * Funding Pressure		0.01 (-0.04, 0.06)		1.00 (0.89, 1.11)
Organised Scepticism * Publication Pressure		-0.01 (-0.06, 0.03)		1.05 (0.95, 1.15)
Communalism * Publication Pressure		0.02 (-0.02, 0.06)		1.02 (0.93, 1.12)
Communalism * Funding Pressure		0.01 (-0.02, 0.05)		1.04 (0.94, 1.15)
Conscientiousness * Publication Pressure		0.03 (-0.01, 0.07)		1.02 (0.93, 1.11)
Neuroticism * Publication Pressure		-0.04 (-0.08, 0.01)		**0.91 (0.83, 1.00)**
Openness * Publication Pressure		0 (-0.04, 0.03)		0.97 (0.88, 1.06)
Agreeableness * Publication Pressure		-0.01 (-0.04, 0.02)		1.02 (0.93, 1.12)

n = 3050; Coefficients and confidence intervals (CIs) are based on pooled estimates across 15 imputed datasets, using Rubin's rules. CIs are bootstrapped. Sampling weights were included in all models.
Overall mean QRP was computed based on 4-12 QRP items with not applicable responses removed.
Any frequent QRP is at least one of the 12 QRPs being often or always being engaged in.
All models contain the five background characteristics of sex, disciplinary field, methodological approach, career stage and country.
Bolded figures indicate statistical significance at the **0.05** alpha level.

Table 4.8. Regression Coefficients (95% confidence interval) of proximal science-specific motivators (achievement motivation and impact and quality motivation)

	Achievement Motivation	Impact and Quality Motivation
	Linear Regression Model 3A Coefficient (95% CI)	**Linear Regression Model 3B Coefficient (95% CI)**
Publication Pressure	**0.08 (0.05, 0.12)**	0.01 (-0.03, 0.05)
Funding Pressure	**0.14 (0.10, 0.18)**	**0.05 (0.00, 0.09)**
Personal Motivations		
Career Growth	**0.34 (0.30, 0.37)**	0.05 (-0.00, 0.09)
Ego and Esteem	**0.25 (0.22, 0.29)**	**0.15 (0.11, 0.19)**
Career Growth * Funding Pressure	-0.02 (-0.06, 0.01)	0.00 (-0.04, 0.05)
Career Growth * Publication Pressure	0.02 (-0.01, 0.06)	0.00 (-0.04, 0.05)
Ego and Esteem * Funding Pressure	0.00 (-0.04, 0.03)	0.00 (-0.04, 0.05)
Ego and Esteem * Publication Pressure	-0.03 (-0.07, 0.01)	-0.04 (-0.09, 0.01)

n = 3050; Coefficients and confidence intervals (CIs) are based on pooled estimates across 15 imputed datasets, using Rubin's rules. CIs are bootstrapped. Sampling weights were included in all models. Models include five background characteristics (sex, disciplinary field, methodological approach, career stage and country) and left-right political orientation scale. Bolded figures indicate statistical significance at the **0.05** alpha level.

Table 4.9. Fit indices for structural models of direct and indirect associations of mean QRP (DV), achievement motivation (M1) and impact and quality motivation (M2)

	Model 4A	Model 4B	Model 4C (Final)
Tucker-Lewis Index (TLI)	-0.72	0.89	0.99
Root Mean Square Error of Approximation (RMSEA)	0.15	0.04	0.01

CFI (Comparative Fit Index)	0.96	0.96	0.99

n = 3088; Fit statistics are based on estimates mean averaged across 15 imputed datasets.
Sampling weights were included in all models.
Model 1 included all five background characteristics (sex, disciplinary field, methodological approach, career stage, and country) as well as the left-right political orientation scale. Model 4B removed non-significant demographic controls ($p > .05$). Model 4C added a covariance between M1 and M2, based on theoretical rationale and modification indices.

The final structural model (Model 4C) yielded an excellent fit to the data (TLI = 0.99, RMSEA = 0.01, CFI = 0.99). Results indicated that there is a statistically significant and positive relationship between the personal motivation of 'ego & esteem' ($\beta = 0.14$, $p < 0.05$) and the proximal science-specific motivation of 'scientific achievement' ($\beta = 0.09$, $p < 0.05$**)** and average self-reported QRP engagement. The proximal science-specific motivation of 'impact and quality' was found to have a statistically significant negative association with average self-reported QRP engagement. That is, as researchers are more motivated to improve the impact and quality of their research, their average self-reported QRP engagement decreases. We find no statistically significant direct association of the personal motivation of career growth on average self-reported QRP engagement, but we do find a statistically significant indirect association through the proximal science-specific motivations of achievement ($\beta = 0.04$, $p < 0.05$) and impact and quality ($\beta = -0.01$, $p < 0.05$). The results additionally indicate that the relationship between the personal motivation of ego and esteem on average self-reported QRP engagement are in part mediated through both proximal science-specific motivations of achievement ($\beta = 0.02$, $p < 0.05$) and impact and quality ($\beta = -0.03$, $p < 0.05$).

Table 4.10. Path coefficients (standard errors) and fit statistics for structural model of direct associations of mean QRP (Y), achievement motivation (M1) and impact and quality motivation (M2)

	SEM Model 4A Standardised Coefficient (SE)	**SEM Model 4B Standardised Coefficient (SE)**	**Model 4C (Final) Standardised Coefficient (SE)**
Career growth motivation (X1) → achievement motivation (M1)	**0.39 (0.02)**	**0.39 (0.02)**	**0.39 (0.02)**
Ego and esteem motivation (X2) → achievement motivation (M1)	**0.26 (0.02)**	**0.26 (0.02)**	**0.26 (0.02)**

Career growth motivation (X1) → Impact and quality motivation (M2)	**0.06 (0.02)**	**0.07 (0.02)**	**0.07 (0.02)**
Ego and esteem motivation (X2) → Impact and quality motivation (M2)	**0.15 (0.02)**	**0.15 (0.02)**	**0.15 (0.02)**
Achievement motivation (M1) → mean QRP (Y)	**0.10 (0.02)**	**0.09 (0.02)**	**0.09 (0.02)**
Impact and quality motivation (M2)→ mean QRP (Y)	**-0.18 (0.02)**	**-0.19 (0.02)**	**-0.19 (0.02)**
Career growth motivation (X1) → mean QRP (Y)	-0.02 (0.02)		
Ego and esteem motivation (X2) → mean QRP (Y)	**0.15 (0.02)**	**0.14 (0.02)**	**0.14 (0.02)**
R^2 Mean QRP	0.09 (0.00)	0.09 (0.00)	0.09 (0.00)
R^2 Achievement Motivation	0.35 (0.00)	0.35 (0.00)	0.35 (0.00)
R^2 Impact and Quality Motivation	0.08 (0.00)	0.08 (0.00)	0.08 (0.00)

n = 3050; Coefficients and SEs are based on pooled estimates across 15 imputed datasets, using Rubin's rules.
Sampling weights were included in all models.
Model 1 included all five background characteristics (sex, disciplinary field, methodological approach, career stage, and country) as well as the left-right political orientation scale. Model 2 removed non-significant demographic controls (p > .05). Model 3 added a covariance between M1 and M2, based on theoretical rationale and modification indices.
Bolded figures indicate statistical significance at the **0.05** alpha level.

Table 4.11. Path coefficients (standard errors) for structural model of indirect associations of career growth motivations and ego and esteem motivations on mean QRP

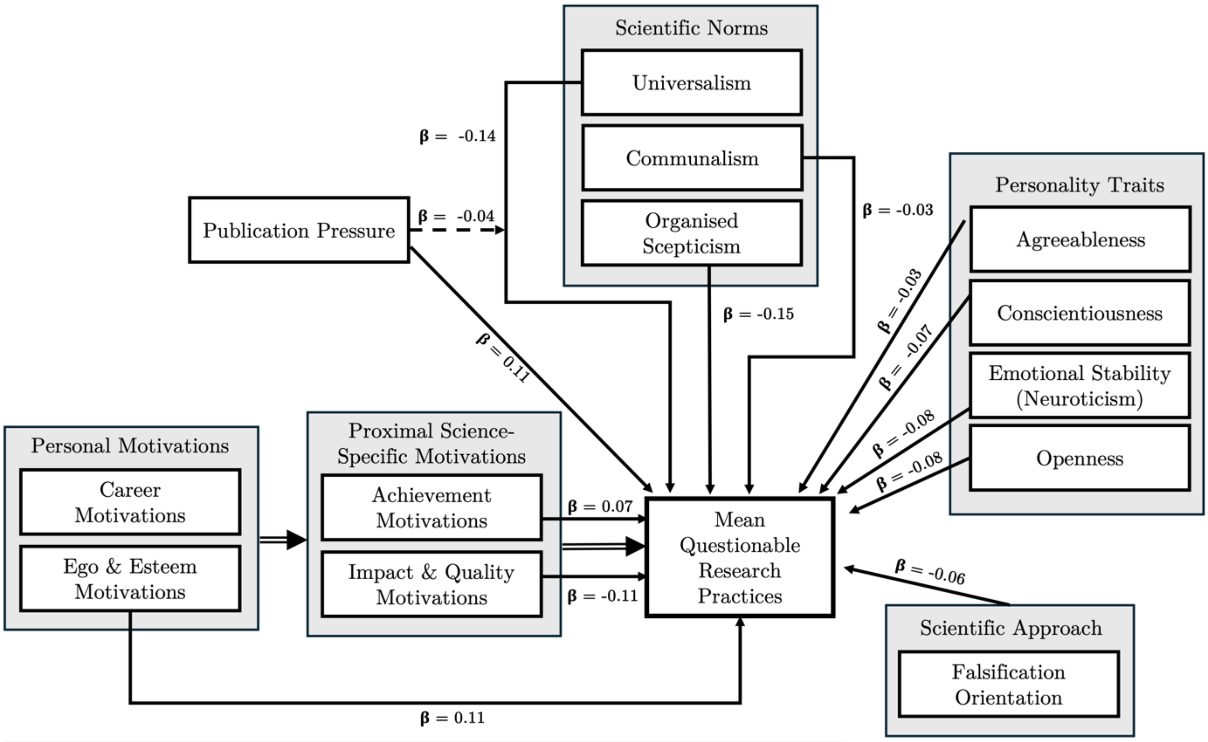
	SEM Model 1 Standardised Coefficient (SE)	SEM Model 2 Standardised Coefficient (SE)	Model 3 (Final) Standardised Coefficient (SE)
Career growth motivation (X1) → achievement motivation (M1) → mean QRP (Y)	**0.04 (0.01)**	**0.04 (0.01)**	**0.04 (0.01)**
Ego and esteem motivation (X2) → achievement motivation (M1) → mean QRP (Y)	**0.03 (0.01)**	**0.02 (0.01)**	**0.02 (0.01)**
Career growth motivation (X1) → Impact and quality motivation (M2) → mean QRP (Y)	**-0.01 (0.00)**	**-0.01 (0.00)**	**-0.01 (0.00)**
Ego and esteem motivation (X2) → Impact and quality motivation (M2) → mean QRP (Y)	**-0.03 (0.00)**	**-0.03 (0.00)**	**-0.03 (0.01)**

n = 3050; Coefficients and SEs are based on pooled estimates across 15 imputed datasets, using Rubin's rules.
Sampling weights were included in all models.
Model 1 included all five background characteristics (sex, disciplinary field, methodological approach, career stage, and country) as well as the left-right political orientation scale. Model 2 removed non-significant demographic controls (p > .05). Model 3 added a covariance between M1 and M2, based on theoretical rationale and modification indices. Bolded figures indicate statistical significance at the **0.05** alpha level.

The Final Conceptual Model

Figure 4.5 presents our final integrated model, which synthesises statistically significant relationships from across separate regression, mediation and moderation analyses. Only predictors that reached statistical significance are retained for representation in this final model, resulting in a parsimonious model of social and psychological correlates of self-reported QRP engagement.

Figure 4.5. Final empirical model illustrating the identified social-psychological correlates of self-reported engagement in questionable research practices



Indirect Paths (Mediation)
Career growth motivation → achievement motivation → mean QRP (0.04)
Ego and esteem motivation → achievement motivation → mean QRP (0.02)
Career growth motivation → Impact and quality motivation → mean QRP (-0.01)
Ego and esteem motivation → Impact and quality motivation → mean QRP (-0.03)

Note. The coefficients shown are synthesised from separate analyses rather than derived from a single integrated model. Direct comparisons of effect sizes should be interpreted with caution. Only statistically significant predictors are displayed in this final model. Non-significant paths from the initial conceptual model were removed to yield a parsimonious representation. The compound (double-lined) arow indicates the indirect paths from personal motivations through proximal science-specific motivations.

Discussion

In this study, we present a comprehensive empirical model of the correlates of self-reported engagement in QRPs, drawing on data from both our Research Integrity and Culture Survey and the International Research Integrity Survey (IRIS). Further details about IRIS and reproducibility materials are contained within Allum et al. (2023). The Research Integrity and Culture Survey is one of the largest cross-national investigations to date into the social, psychological, and cultural factors associated with self-reported QRP engagement, encompassing researchers from across all disciplinary fields in Europe and the United States (see, Xie et al., 2021). Our findings indicate prevalence estimates of up to 60% for certain QRPs, aligning with rates reported in previous large-scale surveys (e.g., Schneider et al., 2024; Gopalakrishna et al., 2022).

Main disciplinary field was not consistently associated with self-reported QRP engagement. The only statistically significant disciplinary difference was observed for respondents in the humanities, who reported lower average QRP engagement than respondents in the natural sciences. This finding should be interpreted cautiously. Lower self-reported engagement in the humanities may not straightforwardly indicate greater research integrity, but may instead reflect disciplinary differences in the applicability, recognition, or normative interpretation of the QRP items. Practices such as selective reporting, HARKING, excluding data, or failing to report null findings may have clearer meanings in experimental, quantitative, or hypothesis-testing traditions than in interpretive, theoretical, historical, or qualitative research. Thus, the observed difference may partly reflect differences in methodological traditions and opportunity structures, rather than differences in willingness to compromise research integrity. More broadly, the limited disciplinary associations suggest that broad field categories may be too coarse to capture the specific normative and institutional contexts that shape QRP engagement.

Similarly, we find that commitment to the Mertonian ideals of universalism, communalism and organised scepticism are widespread, but not universal, consistent with findings elsewhere (Anderson et al., 2010; Bray & Storch, 2017; Macfarlane & Cheng, 2008; Kim & Kim, 2018).

Our research supports pre-eminent theorising that misaligned incentive structures and the resulting systemic pressures, at least in part, motivate engagement in QRPs (Bouter, 2015). This aligns with findings from comparable surveys – often conducted within specific disciplinary fields or national contexts – which consistently identify publication pressure as among the strongest correlates of QRP engagement (Schneider et al., 2024; Gopalakrishna et al., 2022; Maggio et al., 2019; Latan et al., 2021; Necker et al., 2014). These findings underscore the importance of cultivating a research culture that prioritises collegiality and collaboration, while actively mitigating hyper-competition and the excessive emphasis on publication output. We echo calls made elsewhere that science needs to 'slow down'. As the English statistician Douglas Altman (1994) argued, 'we need less research, better research, and research done for the right reasons.' There is an imperative to cultivate a scientific culture that prioritises quality over speed and publication volume, and one that actively challenges incentive structures which reward mass production (Chin, 2025). Such incentives not only compromise the integrity of the research process but also increase the likelihood of error, making both research and researchers more vulnerable to lapses in rigour.

We found no evidence to support a significant association between funding pressure and extent of self-reported engagement in QRPs. One possible explanation is that the behaviours captured under the QRP construct – though broad and extensive – may not directly reflect practices that offer a competitive advantage in securing research funding. As such, the link between funding-related pressures and QRPs may be more nuanced or operate through alternative mechanisms not captured in our research.

Scientific norms emerged as one of the strongest correlates of self-reported QRP engagement. The greater a researcher's commitment to universalism, communalism and organised scepticism, the less likely they were to report engaging in QRPs, with comparably stronger associations for universalism and organised scepticism than for communalism. Further, our findings indicate that a researcher's self-reported subscription to universalism and organised scepticism are perhaps even more influential than publication pressure in determining engagement in QRPs. Our findings here align with prior research demonstrating that stronger commitment to scientific norms – when conceptualised as a unified construct – are associated with lower levels of self-reported engagement in QRPs and scientific misconduct (Brooker & Allum, 2024; Gopalakrishna et al., 2022; Anderson et al., 2010). However, we extend these findings by suggesting that the qualitative characteristics of individual scientific norms matter, with certain normative ideals potentially exerting a stronger influence on QRP engagement than others.

On this basis, we echo recommendations that we have made elsewhere (see, Brooker & Allum, 2024), that institutional leaders should place greater emphasis on fostering commitment to scientific norms as a means of enhancing adherence and promoting high standards of research practice (Martinson et al., 2010; Anderson et al., 2007). On the one hand, this may involve interventions such as normative-affirmation exercises, community pledges, norm-based training, and the cultivation of scientific norms through mentoring and supervisory relationships. On the other hand, greater efficacy may lie in fostering pathways into research careers for individuals who exhibit a strong adherence to Mertonian norms from the outset. Incorporating assessments of 'value fit' into faculty recruitment processes may help institutions ensure alignment between candidates' normative commitments and the organisation's research integrity ethos. It may not be unreasonable to suggest that such efforts are more tangible and feasible than broader systemic-level reforms – particularly as they are more readily actionable by institutional and departmental leaders. Moreover, our results indicate that certain normative ideals, notably universalism, may act as buffers against the effects of systemic pressures, potentially reducing the likelihood of QRP engagement. However, further research is needed to elucidate the mechanisms that underlie these relationships.

We additionally find that epistemic orientation is important in shaping research practice. Researchers who were motivated to falsify their own findings and theories – or those of other researchers – were significantly less likely to engage in QRPs. Conversely, while trending in the expected direction, we find no statistically significant relationship between a researcher's motivation to confirm their own findings and theories – and those of others – and QRP engagement. However, the negative association between falsification and self-reported QRP engagement suggests that actively seeking to test and potentially disconfirm one's assumptions may act as a protective factor. This aligns with longstanding scientific ideals, particularly those rooted in Popperian falsification, which emphasise critical scrutiny and the systematic testing of competing explanations. This complements our finding that researchers with a greater commitment to the Mertonian norm of organised scepticism engage in fewer QRPs on average.

Beyond a researchers 'value-system' and epistemic orientation, we find that their personal and proximal-science specific motivations correlate with engagement in QRPs. This provides empirical validity to the social-psychological model of researcher practice developed by Jussim et al. (2019). Our findings reveal an indirect association between heightened career motivation – such as aspirations for increased salary or promotion – and self-reported QRP engagement, mediated by a stronger drive for scientific achievement, operationalised through frequent publication and citation accumulation. These results suggest that career motivation, in itself, is not inherently problematic; rather, it becomes a risk factor when academic success is defined by metrics that are susceptible to gaming or manipulation. Promoting career advancement frameworks that reward quality-enhancing practices – such as open science, transparency, and methodological integrity – rather than quantity-based output metrics, may offer a viable strategy for reducing engagement in QRPs.

However, an additional indirect pathway suggests a positive influence of career motivation on self-reported QRP engagement, wherein career growth ambitions foster a stronger motivation for conducting impactful and high-quality research. This increased focus on impact and quality is negatively associated with QRP engagement, suggesting that when researchers perceive career advancement as tied to the broader scientific or societal impact of their work, they are less likely to compromise methodological integrity. Although this positive pathway is weaker than the publication-metric-driven pathway, it indicates that institutions that effectively align career advancement with research quality and impact not only discourage QRPs but may actively promote a healthier research culture. Our finding that career motivation that fosters impactful and high-quality research, and subsequently reduced self-reported engagement in QRPs, is consistent with Janke et al.'s (2019) observation that achievement goals oriented toward skill-development are associated with reduced QRP engagement. Conversely, our finding that where career motivation intensifies the pursuit of quantity-driven metrics (e.g., publication and citation counts) there is increased QRP engagement, mirrors Janke et al.'s (2019) finding that researchers focused on skill demonstration – rather than skill development – are more likely to engage in QRPs.

Where researchers were motivated to obtain fame, avoid external perceptions of failure and achieve political or social goals (here termed 'ego and esteem motivations'), average self-reported QRP engagement was higher. These motivations reflect a desire for self-enhancement and external validation which come at cost of epistemic and integrity concerns. On this basis, research cultures should be motivated by intrinsic, epistemic concerns, rather than extrinsic, image-oriented concerns. Interestingly, we did *not* find that political orientation correlates with self-reported QRP engagement. On this basis, it's feasible that a researcher's political orientation (i.e., whether they're left- or right- leaning) per se does not matter, but rather the instrumentalisation of research for ideological or advocacy-related ends *does* matter. When scientific research becomes a means to advance personal, social or political objectives, the risk of compromised integrity appears to increase.

Finally, higher levels of conscientiousness, agreeableness, emotional stability (opposite to neuroticism) and openness among researchers were negatively associated with self-reported QRP engagement. This suggests that researchers who are more diligent, cooperative, emotionally sensitive and open to new experiences are perhaps less inclined to compromise their research integrity and engage in QRPs. This is consistent with earlier research showing that greater conscientiousness, agreeableness, emotional stability and openness may act as protective traits against challenges to research integrity, such as QRPs (e.g., Schneider et al., 2024; Giluk & Postlethwaite, 2015). Schneider et al. (2024) found only minuscule associations between extraversion and self-reported QRP engagement. Similarly, we failed to find a statistically significant relationship between extraversion and self-reported QRP engagement.

Limitations

This sample comprised individuals who previously participated in IRIS and self-selected as willing to be contacted for future research. As a result, certain characteristics may be overrepresented in the data, potentially limiting representativeness. In addition, because QRP engagement is self-reported and socially sensitive, absolute prevalence estimates may be biased downwards due to underreporting. This limitation is particularly salient given the low response rate. However, prior IRIS participation may increase respondents' awareness of what constitutes a QRP and their willingness (or ability) to recognise and report it, which could inflate estimates relative to a less 'integrity-aware' sample. While these issues bear most directly on prevalence, the primary interest is in covariation: non-differential misreporting would mainly attenuate associations, whereas differential misreporting correlated with key predictors could bias estimates in either direction. This limitation is particularly relevant for the personality findings. Traits such as agreeableness, conscientiousness, emotional stability, and openness are also associated with more socially desirable self-presentation, raising the possibility that some of the observed negative associations with self-reported QRP engagement reflect differences in willingness to admit undesirable practices rather than differences in actual behaviour. Because the study did not include a social desirability measure or external behavioural validation, this possibility cannot be ruled out. The personality results should therefore be interpreted as associations with self-reported QRP engagement specifically. Nonetheless, confidence in the substantive patterns is strengthened by their consistency – in direction and magnitude – with findings in the extant literature, and by the inclusion of relevant controls (including methodological orientation) in the models. However, the results are best interpreted as evidence of associations within the sample, rather than estimates of population-level prevalence or effect sizes.

Moreover, the proportion of mixed-methods researchers in our sample far exceeds what would typically be expected in the broader researcher population, suggesting a potential sampling imbalance that may limit representativeness. However, we remain confident in the validity of our results, given the consistency of effect sizes and patterns with those reported in the extant literature. Notably, we account for an over-representation of mixed-method researchers by controlling for methodological orientation in our models. The sample additionally consisted of researchers primarily employed in Europe and North America, so we do not seek to generalise our findings beyond the European and Anglo-American sphere. The overall response rate was low, but within the range of other surveys (Gopalkrishna et al., 2024; Fanelli, 2009; Haven, 2019).

A further limitation is that the effect sizes of all independent variables are small. However, this is consistent with patterns observed in much of the extant literature, where associations with QRP engagement tend to be modest but meaningful. The associations may be attenuated by the limited variability in self-reported QRP engagement and responses. Likewise, small probabilities of engagement may translate into meaningful differences when aggregated across researchers. That being said, we should be humble in our interpretation of the results.

While our analysis includes three of the four classical Mertonian norms – organised scepticism, universalism, and communalism – we do not include disinterestedness. This decision was informed by our factor analysis, which did not identify a coherent or reliable factor corresponding to disinterestedness. Certainly, this Mertonian norm has proven difficult to operationalise and measure reliably, whereby agreement with disinterestedness-affirming statements differs significantly across items (Macfarlane & Cheng, 2008), thus undermining its coherency as a normative construct. Financial, reputational and ideological interests (as demonstrated in our study) are often embedded within the scientific enterprise and undermine support for the normative ideal of disinterestedness.

Finally, although the study identifies social and psychological factors associated with self-reported QRP engagement, the cross-sectional survey design precludes conclusions about causal direction. For example, while publication pressure, career-achievement motivations, esteem-related motivations, and weaker endorsement of Mertonian norms were associated with greater self-reported QRP engagement, these relationships should not be interpreted as demonstrating that such factors *cause* QRP engagement. Reverse or reciprocal interpretations are also plausible. Researchers who report engaging in QRPs may retrospectively justify those practices by perceiving greater pressure, placing greater emphasis on career-related incentives, or reporting weaker commitment to scientific norms. Similarly, self-reported QRP engagement may shape how researchers interpret their research environment and motivations, rather than only being shaped by them.

The observed associations may also reflect non-causal explanations, including unmeasured disciplinary, institutional, or cultural factors that influence both researchers' motivations and their reported engagement in QRPs. In addition, because both predictors and outcomes were measured through self-report, associations may partly reflect shared method variance, differences in self-presentation, recall, or willingness to disclose sensitive behaviours. It is therefore important to emphasise that the outcome captures reported engagement in QRPs rather than directly observed behaviour. Accordingly, the findings should be interpreted as evidence of associations between social and psychological characteristics and self-reported QRP engagement, rather than as evidence that these factors causally increase or reduce QRP engagement.

Conclusion

The prevalence and downstream impact underscore the need for a robust explanatory framework of QRP engagement. Here, we identify a set of individual-level social and psychological variables that meaningfully explain variation in self-reported engagement in these practices. We identify several opportunities to develop leverage points aimed at reducing engagement in QRPs and strengthening research integrity. Foremost, we have shown that individual-level factors, such as a researcher's normative commitment and a researcher's personal and scientific motivations, matter as much as systemic-level pressures such as publication pressure, in increasing QRP engagement. Designing interventions and policies that promote integrity-enhancing ideals, values, and motivations represents a feasible leverage point – one that may be more readily adjustable than broader systemic or structural reforms. That said, it would be unwise to interpret this as evidence that systemic pressures are unimportant; on the contrary, their influence remains evident and significant. Therefore, mitigating the use of QRPs and maximising the validity of research requires a multifaceted approach that targets both systemic structures (i.e., 'publish or perish' culture) and the individual motivations, values, and beliefs that drive research conduct. A multifaceted, evidence-based approach will permit the mitigation of QRP engagement, and therefore improve the quality, validity and reliability of scientific research

References

Allum, N., Reid, A., Bidoglia, M., Gaskell, G., Aubert-Bonn, N., Buljan, I., Fuglsang, S., Horbach, S., Kavouras, P., Marušić, A., Mejlgaard, N., Pizzolato, D., Roje, R., Tijdink, J., & Veltri, G. (2023). Researchers on research integrity: a survey of European and American researchers. *F1000Research*, *12*, 187. https://doi.org/10.12688/f1000research.128733.1

Altman D. G. (1994). The scandal of poor medical research. *BMJ (Clinical research ed.)*, *308*(6924), 283–284. https://doi.org/10.1136/bmj.308.6924.283

Anderson, M. S., & Louis, K. S. (1994). The graduate student experience and subscription to the norms of science. Research in Higher Education, *35*, 273-299. https://doi.org/10.1007/BF02496825

Anderson, M. S., Martinson, B. C., & De Vries, R. (2007). Normative dissonance in science: results from a national survey of u.s. Scientists. *Journal of empirical research on human research ethics : JERHRE*, *2*(4), 3–14. https://doi.org/10.1525/jer.2007.2.4.3

Anderson, M. S., Ronning, E. A., Devries, R., & Martinson, B. C. (2010). Extending the Mertonian Norms: Scientists' Subscription to Norms of Research. *The Journal of higher education*, *81*(3), 366–393. https://doi.org/10.1353/jhe.0.0095

Artino, A. R., Jr, Driessen, E. W., & Maggio, L. A. (2019). Ethical Shades of Gray: International Frequency of Scientific Misconduct and Questionable Research Practices in Health Professions Education. *Academic medicine : journal of the Association of American Medical Colleges*, *94*(1), 76–84. https://doi.org/10.1097/ACM.0000000000002412

Barnes, B. (1982). *T.S. Kuhn and Social Science*. Columbia University Press.

Bouter L. M. (2015). Commentary: Perverse incentives or rotten apples?. *Accountability in research*, *22*(3), 148–161. https://doi.org/10.1080/08989621.2014.950253

Bray, D., & von Storch, H. (2017). The Normative Orientations of Climate Scientists. *Science and engineering ethics*, *23*(5), 1351–1367. https://doi.org/10.1007/s11948-014-9605-1

Brooker, R. (20245, September 16). Individual-level and system-level determinants of questionable research practice. https://doi.org/10.17605/OSF.IO/2BS6U

Brooker, R., & Allum, N. (2024). Investigating the links between questionable research practices, scientific norms and organisational culture. *Research integrity and peer review*, *9*(1), 12. https://doi.org/10.1186/s41073-024-00151-x

Chambers, J. R., & Schlenker, B. R. (2015). Political homogeneity can nurture threats to research validity. *Behavioral and Brain Sciences*, *38*, e138. doi:10.1017/S0140525X14001162

Chin, J. (2025, March 27). *Reliable science takes time. But the current system rewards speed.* The Conversation. https://theconversation.com/reliable-science-takes-time-but-the-current-system-rewards-speed-249497#:~:text=Slow%20science%20may%20not%20represent,And%20this%20takes%20time.

Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., Celniker, J. B., & Zinger, J. F. (2018). At Least Bias Is Bipartisan: A Meta-Analytic Comparison of Partisan Bias in Liberals and Conservatives. *Perspectives on Psychological Science*, *14*(2), 273-291. https://doi.org/10.1177/1745691617746796

Duarte, J. L., Crawford, J. T., Stern, C., Haidt, J., Jussim, L., & Tetlock, P. E. (2015). Political diversity will improve social psychological science. *The Behavioral and brain sciences*, *38*, e130. https://doi.org/10.1017/S0140525X14000430

Edwards, M. A., & Roy, S. (2017). Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition.

Environmental engineering science, *34*(1), 51–61. https://doi.org/10.1089/ees.2016.0223

Fanelli D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PloS one, 4(5), e5738. https://doi.org/10.1371/journal.pone.0005738

Fiedler, K., & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological and Personality Science, 7*(1), 45–52. https://doi.org/10.1177/1948550615612150

Francis, G. (2012) 'Publication bias and the failure of replication in experimental psychology', *Psychonomic Bulletin & Review*, 19(6), pp. 975–991. Available at: https://doi.org/10.3758/s13423-012-0322-y (Accessed: 3 June 2026).

Fraser, H., Parker, T., Nakagawa, S., Barnett, A., & Fidler, F. (2018). Questionable research practices in ecology and evolution. *PloS one*, *13*(7), e0200303. https://doi.org/10.1371/journal.pone.0200303

Fuller, S. (2003). *Kuhn vs. Popper : the struggle for the soul of science.* Cambridge.

Gharbi, M.A. (2018). Race and the Race for the White House: On Social Research in the Age of Trump. *The American Sociologist, 49*, 496-519. https://doi.org/10.1007/s12108-018-9373-5

Giluk, T. L., & Postlethwaite, B. E. (2015). Big five personality and academic dishonesty: A meta-analytic review. *Personality and Individual Differences, 72,* 59–67. https://doi.org/10.1016/j.paid.2014.08.027

Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PloS one*, *17*(2), e0263023. https://doi.org/10.1371/journal.pone.0263023

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Gross, N. (2013). *Why Are Professors Liberal and Why Do Conservatives Care?* Harvard University Press.

Gross, N., & Fosse, E. (2012). Why are professors liberal? *Theory and Society*, *41*(2), 127-168. https://doi.org/10.1007/s11186-012-9163-y

Gross, N., & Simmons, S. (2014). *Professors and their politics*. Baltimore: Johns Hopkins University Press. https://dx.doi.org/10.1353/book.31449.

Haven, T. L., Tijdink, J. K., Martinson, B. C., & Bouter, L. M. (2019). Perceptions of research integrity climate differ between academic ranks and disciplinary fields: Results from a survey among academic researchers in Amsterdam. *PloS one*, *14*(1). https://doi.org/10.1371/journal.pone.0210599

Janke, S., Daumiller, M., & Rudert, S. C. (2019). Dark pathways to achievement in science: Researchers' achievement goals predict engagement in questionable research practices. *Social Psychological and Personality Science, 10*(6), 783–791. https://doi.org/10.1177/1948550618790227

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. https://doi.org/10.1177/0956797611430953

Jussim, L., Krosnick, J. A., Stevens, S. T., & Anglin, S. M. (2019). A Social Psychological Model of Scientific Practices: Explaining Research Practices and Outlining the Potential for Successful Reforms. *Psychologica Belgica*, *59*(1), 353–372. https://doi.org/10.5334/pb.496

Kerr N. L. (1998). HARKing: hypothesizing after the results are known. *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc*, *2*(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4

Kim, S. Y., & Kim, Y. (2018). The Ethos of Science and Its Correlates: An Empirical Analysis of Scientists' Endorsement of Mertonian Norms. *Science, Technology and Society*, *23*(1), 1-24. https://doi.org/10.1177/0971721817744438

Klein, D. B., Strern, C., & Western, A. (2005). Political diversity in six disciplines. *Academic Question, 18*(1), 40-52. 10.1007/s12129-004-1031-4

Kuhn, T. S. (1970). *The Structure of Scientific revolutions* (2nd ed.). University of Chicago Press.

Latan, H., Chiappetta Jabbour, C. J., Lopes de Sousa Jabbour, A. B., & Ali, M. (2023). Crossing the red line? Empirical evidence and useful recommendations on questionable research practices among business scholars. *Journal of Business Ethics, 184*(3), 549–569. https://doi.org/10.1007/s10551-021-04961-7

Macfarlane, B., & Cheng, M. (2008). Communism, Universalism and Disinterestedness: Re-examining Contemporary Support among Academics for Merton's Scientific Norms. *Journal of Academic Ethics*, *6*(1), 67-78. https://doi.org/10.1007/s10805-008-9055-y

Macfarlane, B., Cheng, M. (2008). Communism, Universalism and Disinterestedness: Re-examining Contemporary Support among Academics for Merton's Scientific Norms. *Journal of Academic Ethics* *6*(1), 67–78. https://doi.org/10.1007/s10805-008-9055-y

Maggio, L., Dong, T., Driessen, E., & Artino, A., Jr (2019). Factors associated with scientific misconduct and questionable research practices in health professions education. *Perspectives on medical education*, *8*(2), 74–82. https://doi.org/10.1007/s40037-019-0501-x

Martinson, B. C., Anderson, M. S., & de Vries, R. (2005). Scientists behaving badly. *Nature, 435*(7043), 737–738. https://doi.org/10.1038/435737a

Martinson, B. C., Crain, A. L., De Vries, R., & Anderson, M. S. (2010). The importance of organizational justice in ensuring research integrity. Journal of empirical research on human research ethics : JERHRE, 5(3), 67–83. https://doi.org/10.1525/jer.2010.5.3.67

Merton, R. K. (2010). *Robert K. Merton: Sociology of Science and Sociology as Science* (C. Calhoun, Ed.). Columbia University Press. https://doi.org/10.7312/calh15112

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS biology*, *16*(3), e2004089. https://doi.org/10.1371/journal.pbio.2004089

Necker, S. (2014). Scientific misbehaviour in economics. *Research Policy*, *43*(10), 1747-1759. https://doi.org/10.1016/j.respol.2014.05.002

Oberauer, K. and Lewandowsky, S. (2019) ‘Addressing the theory crisis in psychology’, *Psychonomic Bulletin & Review*, 26(5), pp. 1596–1618. Available at: https://doi.org/10.3758/s13423-019-01645-2 (Accessed: 3 June 2026).

Redding, R. E. (2001). Sociopolitical diversity in psychology: The case for pluralism. *American Psychologist, 56*(3), 205–215. https://doi.org/10.1037/0003-066X.56.3.205

Reinero, D. A., Wills, J. A., Brady, W. J., Mende-Siedlecki, P., Crawford, J. T., & Van Bavel, J. J. (2020). Is the Political Slant of Psychology Research Related to Scientific Replicability? *Perspectives on Psychological Science*, *15*(6), 1310-1328. https://doi.org/10.1177/1745691620924463

Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons Inc., New York. http://dx.doi.org/10.1002/9780470316696

Rubin, M. (2023) ‘Questionable metascience practices’, *Journal of Trial & Error*, 4(1), pp. 5–20. Available at: https://doi.org/10.36850/mr4 (Accessed: 3 June 2026).

Rubin, M. (2025) *A brief review of research that questions the impact of questionable research practices*. PsyArXiv. Available at: https://doi.org/10.31234/osf.io/ah9wb_v3 (Accessed: 3 June 2026).

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In Defense of the Questionable: Defining the Basis of Research Scientists' Engagement in Questionable Research Practices. *Journal of empirical research on human research ethics : JERHRE*, *13*(1), 101–110. https://doi.org/10.1177/1556264617743834

Schneider, J. W., Allum, N., Andersen, J. P., Petersen, M. B., Madsen, E. B., Mejlgaard, N., & Zachariae, R. (2024). Is something rotten in the state of Denmark? Cross-national evidence for widespread involvement but not systematic use of questionable research practices across all fields of research. *PloS one*, *19*(8), e0304342. https://doi.org/10.1371/journal.pone.0304342

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. https://doi.org/10.1177/0956797611417632

THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH. (2023). *Standardd Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys*. https://aapor.org/wp-content/uploads/2023/05/Standards-Definitions-10th-edition.pdf

Ulrich, R. and Miller, J. (2020) 'Questionable research practices may have little effect on replicability', *eLife*, 9, e58237. Available at: https://doi.org/10.7554/eLife.58237 (Accessed: 3 June 2026).

van Buuren, S., & Groothuis-Oudshoorn, C. G. M. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of statistical software*, *45*(3). http://www.jstatsoft.org/v45/i03

van de Werfhorst H. G. (2020). Are universities left-wing bastions? The political orientation of professors, professionals, and managers in Europe. *The British journal of sociology*, *71*(1), 47–73. https://doi.org/10.1111/1468-4446.12716

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid *p*-hacking. *Frontiers in Psychology, 7,* Article 1832. https://doi.org/10.3389/fpsyg.2016.01832

Xie, Y., Wang, K., & Kong, Y. (2021). Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis. *Science and engineering ethics*, *27*(4), 41. https://doi.org/10.1007/s11948-021-00314-9\

Yarkoni, T. (2018, October 2). *No, it's not The Incentives-it's you.* Talyarkoni. https://talyarkoni.org/blog/2018/10/02/no-its-not-the-incentives-its-you/

Supplementary Material

Table 4.S1. Mean scores for publication and funding pressure items, stratified by disciplinary field, career stage and sex.

Item Wording	Response Scale	Disciplinary Field				Career Stage				Sex		Full Sample
Publication Pressure		**Social Sciences**	**Natural Sciences**	**Medical Sciences**	**Arts and Humanities**	**Early Career**	**Mid-Career**	**Late-career**	**Retired**	**Male**	**Female**	
On a scale from 'very low pressure' to 'very high pressure', how much pressure do you feel there is to publish in your field?	1 (Very Low Pressure) - 7 (Very High Pressure)	5.66	5.6	5.34	5.48	5.79	5.65	5.46	5.05	5.4	5.78	5.56
How would you rate your level of stress concerning your colleagues' evaluation of your publication output, on a scale from 'extremely stressful' to 'not stressful at all'? [pubpress2] [Reverse Coded]	1 (Extremely Stressful) - 7 (Not Stressful At All)	3.85	3.75	3.68	3.95	4.22	3.94	3.52	3.26	3.64	4.03	3.8
To what extent do you agree that your colleagues judge you on the basis of your publications, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	4.84	4.81	4.61	4.85	4.73	4.74	4.91	4.83	4.76	4.84	4.8
To what extent do you agree publication pressure harms your ability to do good research, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	4.11	4.13	3.77	4.09	4.64	4.25	3.75	3.11	3.92	4.25	4.06
To what extent do you agree that the current publication climate puts pressure on your relationships with fellow researchers, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	4.01	4.06	3.89	3.99	4.34	4.17	3.71	3.56	3.88	4.18	4.01
To what extent do you agree that publication pressure sometimes leads you to cut corners, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	2.64	2.62	2.42	2.54	3.06	2.73	2.3	1.86	2.6	2.57	2.58
Funding Pressure												
To what extent do you agree that your ability to obtain research funding depends on the number of publications you have, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	5.45	5.65	5.47	5.35	5.8	5.54	5.44	5.03	5.44	5.59	5.51
To what extent do you agree that judgements of your academic performance are independent of your successful grant applications, on a scale from 'totally disagree' to 'totally agree'? [Reverse Coded]	1 (Totally Disagree) - 7 (Totally Agree)	4.55	4.77	4.59	4.37	4.61	4.72	4.56	4.24	4.56	4.66	4.6

To what extent do you agree your job security depends strongly on research grants you obtain, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	3.57	4.02	3.79	3.55	4.88	3.85	3.06	3.25	3.57	4.01	3.76
To what extent do you agree that your prospects for promotion depend on you obtaining funding, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	4.69	5.19	4.69	4.69	5.37	5.14	4.42	4.18	4.76	5.03	4.88
To what extent do you agree that you would be able to do your research without obtaining your own funding, on a scale from 'totally disagree' to 'totally agree'? [Reverse Coded]	1 (Totally Disagree) - 7 (Totally Agree)	3.89	4.67	4.56	3.63	4.58	4.3	4.16	3.48	4.1	4.4	4.23
To what extent do you agree that obtaining your own research funding is crucial for your academic career, on a scale from 'totally disagree' to 'totally agree'?	1 (Totally Disagree) - 7 (Totally Agree)	5.01	5.58	5.23	4.85	5.8	5.36	4.97	4.32	5.07	5.43	5.23

Table 4.S2. Exploratory Factor Analysis full pattern matrix for scientific norm items

Item	Factor 1 Universalism	Factor 2 Communalism	Factor 3 Organised Scepticism	Communalities
The personal background and characteristics of a researcher should be considered when assessing the validity of their scientific claims. (Reversed)	**0.666**	0.008	0.001	0.444
The country where a researcher is based should influence a journal editor's decision to publish the researcher's work in their journal. (Reversed)	**0.451**	-0.004	0.105	0.234
Researchers should take into account the authors' ideological positions when evaluating the validity of the findings or conclusions presented in a scientific article. (Reversed)	**0.605**	-0.018	-0.079	0.354
The personal beliefs and values of researchers should play a role in the interpretation of data or outcomes in research. (Reversed)	**0.361**	0.039	0.121	0.171
Researchers should aim to consider all new evidence, hypotheses, and theories, even those that challenge or contradict their own work.	-0.007	-0.010	**0.679**	0.457
Researchers should reject scientific findings solely because they contradict well-established knowledge. (Reversed)	0.093	-0.041	**0.368**	0.152
Researchers should thoroughly scrutinise the methodology and inferences of a published study before making judgments about the validity and reliability of its findings.	-0.038	0.039	**0.483**	0.238
Research findings should be a communal resource and not proprietary knowledge owned by individuals or institutions.	-0.002	**0.406**	0.103	0.192
Researchers should be obligated to upload data to public databases and repositories.	-0.014	**0.715**	-0.042	0.503
Researchers should withhold raw data and methodologies to prevent others from duplicating or challenging their work. (Reversed)	0.113	**0.372**	0.192	0.241
Eigenvalue	1.176	0.908	0.815	
% Variance	0.118	0.091	0.082	

Factor loadings above 0.30 are emboldened for clarity.
EFA was conducted using maximum likelihood with oblique (oblimin) rotation.
Communalities (h^2) and eigenvalues for each factor are provided to indicate the proportion of variance explained.

Table 4.S3. Mean scores for personal, proximal-scientific and scientific approach items stratified by disciplinary field, career stage and sex.

Item Wording	Disciplinary Field				Career Stage				Sex		Full Sample
Personal Motivations	**Social Sciences**	**Natural Sciences**	**Medical Sciences**	**Arts and Humanities**	**Early-Career**	**Mid-Career**	**Late-career**	**Retired**	**Male**	**Female**	
Advancing my career or achieving promotion, compared to other personal and professional motivations	3.5	3.43	3.59	3.54	3.77	3.55	3.48	2.81	3.41	3.6	3.49
Securing financial benefits (e.g., an increase in pay) or rewards for myself, compared to other personal and professional motivations	3.3	3.14	3.14	3.28	3.63	3.31	3.11	2.32	3.1	3.37	3.21
Ensuring my job security or reducing the risk of job loss, compared to other personal and professional motivations	3.73	3.66	3.44	3.83	4.4	3.92	3.31	2.4	3.49	3.94	3.68
Obtaining respect or fame, compared to other personal and professional motivations	2.91	2.81	2.95	3.03	2.62	2.85	3.06	3.17	2.91	2.89	2.9
Avoiding being perceived as a failure or unsuccessful in my field, compared to other personal and professional motivations	3.4	3.42	3.23	3.5	3.52	3.4	3.41	3.12	3.28	3.57	3.4
Proximal Science-Specific Motivations											
Obtaining a higher citation index	3.03	3.15	3.34	2.47	3.2	3.11	2.9	2.67	3	3.05	3.02
Increasing my number of publications	4.06	3.9	4	3.81	4.15	4	3.93	3.41	3.9	4.02	3.95
Increasing my opportunities for research funding	3.78	4.12	4	3.7	4.31	4.07	3.86	2.76	3.75	4.15	3.92
Producing new and exciting research findings	5.15	5.35	5.15	5.28	5.18	5.25	5.33	5.13	5.22	5.28	5.25
Maximising the impact of my research outside of academia Clarification: by 'impact outside of academia' we mean effecting change or benefit to society, culture, the economy, public policy or services, health, the environment or quality of life.	4.59	4.22	4.67	4.22	4.45	4.53	4.3	4.05	4.18	4.69	4.4
Ensuring the reproducibility and accuracy of my research findings	4.25	4.71	4.75	4.25	4.36	4.52	4.53	4.54	4.44	4.57	4.49
Scientific Approach											
Confirming my past findings	2.91	3.45	3.54	3.2	3.26	3.15	3.29	3.52	3.19	3.33	3.25
Disconfirming my past findings	2.33	2.58	2.76	2.43	2.48	2.47	2.51	2.64	2.58	2.39	2.5
Disconfirming my own theories	2.47	2.71	2.98	2.51	2.59	2.59	2.68	2.8	2.72	2.52	2.64
Disconfirming competing theories or findings	3.16	3.31	3.41	3.47	3.31	3.21	3.37	3.52	3.41	3.17	3.31

Table 4.S4. Exploratory Factor Analysis full pattern matrix for personal, proximal-scientific and scientific approach

	Factor 1 Achievement Motivation (Proximal Science-Specific)	Factor 2 Falsification Orientation (Approach)	Factor 3 Career Growth Motivation (Personal)	Factor 4 Confirmation Orientation (Approach)	Factor 5 Ego & Esteem Motivation (Personal)	Factor 6 Impact & Quality Motivation (Proximal Science-Specific)	Communalities
Obtaining a higher citation index	**0.684**	0.000	0.022	0.101	0.073	-0.063	**0.548**
Increasing my number of publications	**0.855**	-0.007	0.000	-0.001	-0.030	-0.067	**0.685**
Increasing my likelihood of publication in a prestigious outlet	**0.764**	-0.015	0.018	-0.006	0.028	0.106	**0.673**
Increasing my opportunities for research funding	**0.434**	0.033	0.275	-0.042	-0.061	0.227	**0.471**
Disconfirming my past findings	-0.048	**0.776**	0.035	0.149	-0.014	-0.011	**0.761**
Disconfirming my own theories	0.007	**1.030**	-0.008	-0.069	-0.010	0.003	**0.972**
Disconfirming competing theories or findings	0.077	**0.446**	-0.043	0.133	0.119	0.061	**0.341**
Advancing my career or achieving promotion, compared to other personal and professional motivations	0.125	-0.003	**0.638**	-0.008	0.101	0.041	**0.563**
Securing financial benefits (e.g., an increase in pay) or rewards for myself, compared to other personal and professional motivations	-0.026	0.013	**0.842**	0.020	-0.007	-0.067	**0.663**
Ensuring my job security or reducing the risk of job loss, compared to other personal and professional motivations	0.016	-0.031	**0.566**	0.006	-0.069	0.043	**0.346**
Confirming my past findings	-0.005	-0.009	0.028	**0.998**	-0.004	-0.005	**0.995**
Confirming my own theories	0.122	0.202	-0.054	**0.535**	0.064	0.060	**0.522**
Achieving my social or political goals, compared to other personal and professional motivations	-0.110	-0.016	0.214	-0.021	**0.346**	0.118	**0.183**
Obtaining respect or fame, compared to other personal and professional motivations	0.006	-0.003	-0.003	0.000	**0.999**	-0.001	**0.995**
Avoiding being perceived as a failure or unsuccessful in my field, compared to other personal and professional motivations	0.206	0.015	0.099	-0.025	**0.315**	0.037	**0.245**
Producing new and exciting research findings	0.046	-0.014	-0.059	-0.010	0.019	**0.621**	**0.413**
Maximising the impact of my research outside of academia	-0.096	0.027	0.156	-0.001	0.049	**0.420**	**0.204**

Ensuring the reproducibility and accuracy of my research findings	-0.037	0.070	-0.046	0.245	-0.037	**0.479**	**0.359**
Eigenvalue	**2.079**	**1.912**	**1.606**	**1.401**	**1.264**	**0.895**	
% Variance	**0.115**	**0.106**	**0.089**	**0.078**	**0.070**	**0.050**	

Factor loadings above 0.30 are emboldened for clarity.
EFA was conducted using maximum likelihood with oblique (oblimin) rotation.
Communalities (h^2) and eigenvalues for each factor are provided to indicate the proportion of variance explained.

Table 4.S5. Proportion of respondents selecting each answer category for each QRP item

	Item Wording	Methodological Group	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Always (%)
QRP 1	Choosing references for citation in a way that primarily supports my own findings or beliefs.	Non-empirical	43.4 (41.4, 45.4)	27.3 (25.5, 29.1)	21.5 (19.8, 23.2)	6.6 (5.6, 7.6)	1.2 (0.7, 1.6)
QRP 2	Including individuals as authors on a paper who may not have contributed significantly to the research.	Non-empirical	43.8 (41.8, 45.9)	27.3 (25.5, 29.1)	20.6 (18.9, 22.2)	6.8 (5.8, 7.9)	1.5 (0.9, 2)
QRP 3	Citing literature without having read it.	Non-empirical	45.3 (43.3, 47.3)	35.3 (33.4, 37.2)	15.9 (14.4, 17.3)	3.3 (2.6, 4)	0.3 (0.1, 0.5)
QRP 4	Making adjustments to the analysis, inferences or conclusions in a manuscript that you do not believe are valid in order to satisfy reviewers.	Non-empirical	68.6 (66.7, 70.5)	21.4 (19.7, 23.1)	7.9 (6.9, 9)	1.7 (1.2, 2.3)	0.4 (0.1, 0.6)
QRP 5	Omitting citations of publications that present differing views or findings from one's own work.	Non-empirical	78.3 (76.6, 79.9)	18 (16.4, 19.5)	3.2 (2.5, 4)	0.4 (0.1, 0.6)	0.2 (0, 0.3)
QRP 6	Using content from papers under review for personal research or publication.	Non-empirical	88.2 (86.8, 89.5)	7.5 (6.4, 8.6)	3.3 (2.5, 4)	0.8 (0.5, 1.2)	0.2 (0, 0.4)
QRP 7	Not disclosing relevant conflicts of interest.	Non-empirical	96 (95.1, 96.8)	2.6 (2, 3.3)	0.8 (0.4, 1.2)	0.3 (0.1, 0.6)	0.2 (0, 0.5)
QRP 8	Excluding junior researchers from the list of authors on a paper despite them having made significant contributions.	Non-empirical	98.5 (98, 99)	1 (0.6, 1.4)	0.4 (0.1, 0.6)	0.2 (0, 0.4)	0 (0, 0)
QRP 9	Neglecting to conduct a power analysis to determine the necessary sample size based on the intended significance level, statistical power, and anticipated effect size.	Quantitative	40 (37.1, 42.9)	26.1 (23.5, 28.7)	19.1 (16.8, 21.4)	10.7 (8.9, 12.6)	4.1 (2.9, 5.2)
QRP 10	Not fully disclosing and making accessible all underlying data, computer codes, or syntaxes used in research on open science platforms.	Quantitative	44.4 (41.7, 47.1)	20.2 (18, 22.4)	19.7 (17.5, 21.9)	11.8 (10, 13.5)	3.9 (2.9, 5)
QRP 11	Reporting a set of results as the complete set of analyses when other relevant analyses were also conducted.	Quantitative	59.7 (57.2, 62.3)	22.4 (20.3, 24.6)	12 (10.3, 13.7)	4 (3, 5)	1.9 (1.2, 2.6)
QRP 12	Presenting an unexpected finding or exploratory analysis in a research report as if it were an a priori hypothesis.	Quantitative	66.2 (63.7, 68.8)	18.5 (16.4, 20.5)	10.9 (9.3, 12.6)	2.8 (1.9, 3.7)	1.5 (0.9, 2.2)
QRP 13	Deciding not to report covariates that failed to reach statistical significance (e.g., $p < 0.05$) or some other desired statistical threshold.	Quantitative	66.5 (63.9, 69.1)	19.4 (17.2, 21.6)	10 (8.3, 11.7)	2.5 (1.7, 3.4)	1.6 (0.9, 2.3)
QRP 14	Deciding to exclude data points after first checking the impact on statistical significance (e.g., $p < 0.05$) or some other desired statistical threshold.	Quantitative	78.4 (76.3, 80.6)	13.4 (11.6, 15.2)	6.3 (5, 7.6)	1.2 (0.6, 1.7)	0.7 (0.3, 1.2)
QRP 15	Rounding off p-values so they appear to be below or equal to the desired alpha (e.g., reporting $p = 0.054$ as $p = 0.05$).	Quantitative	87.1 (85.3, 88.9)	6.8 (5.5, 8.2)	4.2 (3.1, 5.2)	1.6 (0.9, 2.3)	0.3 (0, 0.6)
QRP 16	Extending data collection or manipulating data selection during an analysis with the specific intent of obtaining statistically significant results where none were previously found.	Quantitative	87.6 (85.9, 89.3)	7.7 (6.3, 9.1)	3.9 (2.9, 4.9)	0.6 (0.2, 0.9)	0.2 (0, 0.5)
QRP 17	Not maintaining comprehensive and detailed documentation of all decisions, modifications, and	Qualitative	43.8 (40.8, 46.9)	31.6 (28.8, 34.5)	18 (15.6, 20.3)	5.8 (4.4, 7.3)	0.8 (0.2, 1.3)

	steps undertaken throughout the research process.						
QRP 18	Not engaging in reflexivity, which involves failing to acknowledge or address my biases, perspectives, and potential influence on the research process.	Qualitative	52.5 (49.5, 55.6)	35 (32.1, 37.9)	10.1 (8.3, 12)	1.8 (1, 2.6)	0.6 (0.1, 1.1)
QRP 19	Failing to sufficiently describe the methods of data collection and analysis, making it difficult for others to understand how conclusions were reached.	Qualitative	66.4 (63.6, 69.2)	24 (21.5, 26.5)	8.2 (6.5, 9.8)	1.2 (0.6, 1.9)	0.2 (0, 0.5)
QRP 20	Phrasing questions in interviews or focus groups in a way that might influence participants towards certain responses.	Qualitative	66.7 (63.4, 69.9)	25.6 (22.6, 28.6)	6.6 (4.9, 8.3)	0.9 (0.2, 1.5)	0.3 (0, 0.6)
QRP 21	Interpreting qualitative data in a way that confirms the analysts pre-existing ideas and beliefs, ignoring data that might suggest alternative interpretations.	Qualitative	75.4 (72.7, 78.1)	20.4 (17.9, 22.9)	3.7 (2.5, 4.9)	0.2 (0, 0.5)	0.3 (0, 0.6)
QRP 22	Deliberately omitting or failing to discuss data that contradicts or does not align with the emerging themes or patterns identified in the research.	Qualitative	80.2 (77.8, 82.7)	16 (13.8, 18.3)	3.1 (2, 4.1)	0.6 (0.1, 1.1)	0.1 (0, 0.3)
QRP 23	Presenting interview or focus group findings in a manner that may not fully reflect the original data.	Qualitative	81.3 (78.7, 84)	14.4 (12, 16.8)	2.8 (1.7, 4)	1 (0.3, 1.7)	0.4 (0, 0.9)
QRP 24	Providing false or misleading information to participants about the purpose or nature of the study.	Qualitative	95.3 (93.9, 96.6)	2.8 (1.8, 3.9)	1.3 (0.6, 2.1)	0.3 (0, 0.7)	0.2 (0, 0.5

Notes. Percentages are weighted; 95% Wald confidence intervals in brackets, computed with a simple weighted design.

Table 4.S6. Regression Coefficients (95% confidence interval) for individual QRP items

	QRP 1	QRP 2	QRP 3	QRP 4	QRP 5	QRP 6	QRP 7	QRP 8
Item Wording	**Choosing references for citation in a way that primarily supports my own findings or beliefs.**	**Including individuals as authors on a paper who may not have contributed significantly to the research.**	**Citing literature without having read it.**	**Making adjustments to the analysis, inferences or conclusions in a manuscript that you do not believe are valid in order to satisfy reviewers.**	**Omitting citations of publications that present differing views or findings from one's own work.**	**Using content from papers under review for personal research or publication.**	**Not disclosing relevant conflicts of interest.**	**Excluding junior researchers from the list of authors on a paper despite them having made significant contributions.**
Publication Pressure	**0.12 (0.05, 0.18)**	0.06 (-0.01, 0.13)	**0.07 (0.02, 0.13)**	**0.08 (0.04, 0.13)**	0.03 (-0.01, 0.07)	**0.04 (0.00, 0.07)**	0.00 (-0.02, 0.02)	0.00 (-0.01, 0.01)
Funding Pressure	-0.04 (-0.11, 0.03)	0.02 (-0.05, 0.09)	-0.04 (-0.09, 0.02)	0.04 (-0.02, 0.09)	-0.01 (-0.05, 0.02)	**-0.04 (-0.08, -0.01)**	0.00 (-0.02, 0.03)	-0.01 (-0.02, 0.01)
Scientific Norms								
Universalism	-0.13 (-0.20, -0.07)	-0.05 (-0.11, 0.01)	-0.02 (-0.06, 0.03)	0.02 (-0.02, 0.07)	-0.03 (-0.07, 0.00)	**-0.04 (-0.07, -0.01)**	**-0.02 (-0.04, 0.00)**	-0.01 (-0.02, 0.01)
Communalism	-0.03 (-0.09, 0.04)	0.00 (-0.06, 0.06)	0.05 (0.00, 0.10)	-0.01 (-0.06, 0.03)	**-0.02 (-0.06, 0.01)**	0.02 (-0.01, 0.04)	0.01 (-0.02, 0.03)	0.00 (-0.02, 0.02)
Organised Scepticism	**-0.10 (-0.17, -0.03)**	-0.05 (-0.11, 0.01)	-0.04 (-0.10, 0.01)	**-0.07 (-0.12, -0.01)**	-0.05 (-0.08, -0.01)	-0.03 (-0.07, 0.00)	-0.03 (-0.07, 0.01)	-0.02 (-0.05, 0.01)
Personal Motivations								
Career Growth	**-0.10 (-0.16, -0.03)**	-0.02 (-0.09, 0.05)	-0.02 (-0.08, 0.04)	0.00 (-0.05, 0.04)	-0.02 (-0.06, 0.02)	-0.01 (-0.04, 0.02)	0.00 (-0.02, 0.02)	0.01 (-0.01, 0.02)
Ego and Esteem	**0.07 (0.01, 0.14)**	-0.01 (-0.09, 0.06)	**0.08 (0.03, 0.13)**	0.05 (-0.01, 0.10)	**0.05 (0.01, 0.08)**	0.02 (-0.01, 0.06)	0.00 (-0.03, 0.03)	**0.02 (0.00, 0.03)**

Proximal Science-Specific Motivations								
Impact and Quality	-0.05 (-0.13, 0.02)	0.04 (-0.03, 0.11)	**-0.13 (-0.19, -0.07)**	-0.06 (-0.12, -0.01)	**-0.06 (-0.11, -0.01)**	-0.02 (-0.06, 0.01)	-0.01 (-0.06, 0.03)	0.00 (-0.03, 0.03)
Academic Achievement	**0.14 (0.06, 0.22**	0.05 (-0.02, 0.13)	0.00 (-0.06, 0.07	0.03 (-0.03, 0.09)	-0.01 (-0.05, 0.02)	0.00 (-0.04, 0.03)	-0.01 (-0.04, 0.03)	0.00 (-0.01, 0.01)
Research Approach								
Falsification	-0.05 (-0.12, 0.02)	-0.03 (-0.10, 0.04)	**-0.06 (-0.12, 0.00)**	-0.03 (-0.09, 0.02)	-0.02 (-0.06, 0.02)	-0.03 (-0.07, 0.00)	-0.03 (-0.07, 0.01)	0.01 (-0.01, 0.02)
Confirmation	0.08 (0.00, 0.15)	0.02 (-0.06, 0.09)	0.00 (-0.06, 0.06	0.05 (0.00, 0.11)	0.01 (-0.03, 0.06)	0.03 (-0.01, 0.07)	0.04 (0.00, 0.08)	-0.01 (-0.02, 0.01)
Personality Traits								
Extraversion	0.02 (-0.03, 0.08)	**0.07 (0.02, 0.13)**	0.01 (-0.04, 0.06)	**0.04 (0.00, 0.09)**	0.03 (0.00, 0.06)	0.01 (-0.02, 0.04)	0.02 (0.00, 0.03)	0.01 (0.00, 0.02)
Agreeableness	-0.01 (-0.07, 0.04)	0.03 (-0.03, 0.09)	0.01 (-0.04, 0.05)	**-0.06 (-0.11, -0.02)**	-0.02 (-0.05 , 0.01)	0.00 (-0.03, 0.03)	-0.01 (-0.02, 0.01)	0.01 (-0.01, 0.02)
Conscientiousness	-0.01 (-0.06, 0.05)	-0.04 (-0.10, 0.02)	-0.05 (-0.10, 0.00)	-0.03 (-0.08, 0.01)	-0.02 (-0.05, 0.02)	0.01 (-0.02, 0.03)	-0.01 (-0.03, 0.01)	-0.01 (-0.02, 0.01)
Neuroticism	-0.01 (-0.07, 0.05)	**-0.07 (-0.13, -0.01)**	**-0.07 (-0.12, -0.03)**	-0.04 (-0.09, 0.01)	-0.01 (-0.04, 0.02)	0.00 (-0.03, 0.03)	-0.01 (-0.03, 0.01)	-0.01 (-0.02, 0.00)
Openness	**-0.07 (-0.13, -0.01)**	**-0.07 (-0.13, -0.01)**	-0.05 (-0.10, 0.00	**-0.05 (-0.10, 0.00)**	-0.01 (-0.04 , 0.02)	0.01 (-0.02, 0.03)	-0.02 (-0.04, 0.00)	-0.01 (-0.03, 0.00)
Political Orientation	-0.06 (-0.12, 0.00	-0.06 (-0.11, 0.00)	**-0.05 (-0.10, -0.01**	**-0.05 (-0.09, -0.01)**	**-0.04 (-0.07, -0.01)**	-0.01 (-0.04, 0.01	0.02 (-0.01, 0.04)	**0.02 (0.00, 0.03)**

	QRP 9	QRP 10	QRP 11	QRP 12	QRP 13	QRP 14	QRP 15	QRP 16
Item Wording	**Neglecting to conduct a power analysis to determine the necessary sample size based on the intended significance level, statistical power, and anticipated effect size.**	**Not fully disclosing and making accessible all underlying data, computer codes, or syntaxes used in research on open science platforms.**	**Reporting a set of results as the complete set of analyses when other relevant analyses were also conducted.**	**Presenting an unexpected finding or exploratory analysis in a research report as if it were an a priori hypothesis.**	**Deciding not to report covariates that failed to reach statistical significance (e.g., p < 0.05) or some other desired statistical threshold.**	**Deciding to exclude data points after first checking the impact on statistical significance (e.g., p < 0.05) or some other desired statistical threshold.**	**Rounding off p-values so they appear to be below or equal to the desired alpha (e.g., reporting p = 0.054 as p = 0.05).**	**Extending data collection or manipulating data selection during an analysis with the specific intent of obtaining statistically significant results where none were previously found.**
Publication Pressure	0.07 (-0.03, 0.16)	0.00 (-0.09, 0.10)	0.07 (0.00, 0.15)	0.07 (0.00, 0.14)	0.03 (-0.05, 0.10)	**0.06 (0.01, 0.11)**	0.03 (-0.03, 0.08)	0.01 (-0.03, 0.05)
Funding Pressure	0.04 (-0.07, 0.14)	-0.02 (-0.12, 0.08)	-0.03 (-0.10, 0.05)	-0.03 (-0.09, 0.04)	-0.04 (-0.12, 0.04)	-0.05 (-0.10, 0.00)	-0.01 (- 0.07, 0.04)	0.03 (-0.01, 0.07)
Scientific Norms								
Universalism	-0.05 (-0.14, 0.04)	-0.05 (-0.15, 0.04)	-0.02 (-0.09, 0.05)	-0.05 (-0.12, 0.02)	-0.01 (-0.08, 0.06)	**-0.08 (-0.13, -0.02)**	-0.03 (- 0.08, 0.02)	**-0.04 (-0.08, 0.00)**
Communalism	0.02 (-0.08, 0.13)	**-0.13 (-0.23, -0.03)**	-0.02 (-0.09, 0.05)	-0.05 (-0.13, 0.03)	0.02 (-0.06, 0.10)	-0.02 (-0.07, 0.03)	-0.04 (- 0.10, 0.02)	0.00 (-0.04, 0.04)
Organised Scepticism	-0.02 (-0.12, 0.09)	0.04 (-0.07, 0.14)	-0.07 (-0.14, 0.00)	**-0.10 (-0.18, -0.03)**	**-0.11 (-0.20, -0.02)**	**-0.11 (-0.18, -0.03)**	-0.06 (- 0.13, 0.02)	-0.05 (-0.10, 0.00)
Personal Motivations								
Career Growth	0.03 (-0.09, 0.14)	0.02 (-0.07, 0.12)	0.03 (-0.03, 0.10)	0.05 (-0.02, 0.12)	-0.01 (-0.09, 0.07)	0.04 (-0.01, 0.09)	0.01 (-0.04, 0.05)	0.00 (-0.04, 0.03
Ego and Esteem	0.06 (-0.05, 0.16)	**0.10 (0.01, 0.19)**	-0.01 (-0.09, 0.06)	-0.01 (-0.08, 0.06)	0.04 (-0.03, 0.11	0.01 (-0.04, 0.07)	0.03 (-0.02, 0.08	0.00 (-0.04, 0.04)
Proximal Science-Specific Motivations								

Impact and Quality Academic Achievement	-0.08 (-0.19, 0.02) -0.08 (-0.21, 0.06)	0.02 (-0.08, 0.12) 0.04 (-0.06, 0.15)	**-0.14 (-0.21, -0.07)** **0.09 (0.02, 0.16)**	-0.05 (-0.11, 0.01) 0.03 (-0.05, 0.12)	-0.01 (-0.09, 0.07) 0.02 (-0.06, 0.11)	-0.01 (-0.06, 0.04) **0.07 (0.02, 0.13)**	0.00 (-0.05, 0.05) 0.02 (-0.03, 0.07)	-0.03 (-0.07, 0.01) **0.07 (0.03, 0.10**
Research Approach Falsification Confirmation	**-0.13 (-0.24, -0.02)** 0.02 (-0.10, 0.14)	0.01 (-0.10, 0.11) -0.10 (-0.22, 0.01)	**-0.08 (-0.16, -0.01)** **0.08 (0.00, 0.16)**	**-0.05 (-0.13, 0.02)** 0.07 (-0.01, 0.15)	**-0.12 (-0.21, -0.04)** **0.09 (0.00, 0.17**	-0.05 (-0.11, 0.00) 0.01 (-0.05, 0.06)	0.01 (-0.04, 0.05) 0.02 (-0.02, 0.07	**-0.05 (-0.10, -0.01)** 0.03 (-0.01, 0.08)
Personality Traits Extraversion Agreeableness Conscientiousness Neuroticism Openness	0.01 (-0.08, 0.10) -0.04 (-0.14, 0.05) -0.06 (-0.15, 0.03) -0.09 (-0.18, 0.01) -0.11 (-0.2, -0.01)	**-0.09 (-0.18, -0.01)** 0.02 (-0.07, 0.10) -0.01 (-0.09, 0.07) -0.06 (-0.15, 0.02) **-0.15 (-0.24, -0.06)**	0.04 (-0.02, 0.10) -0.06 (-0.12, 0.00) -0.03 (-0.09, 0.03) -0.03 (-0.10, 0.03) **-0.03 (-0.09, 0.04)**	**0.11 (0.05, 0.17)** **-0.07 (-0.13, -0.01)** -0.02 (-0.08, 0.04) 0.00 (-0.06, 0.05) -0.04 (-0.10, 0.02)	-0.01 (-0.07, 0.06) -0.02 (-0.08, 0.04) 0.02 (-0.04, 0.08) -0.05 (-0.13, 0.02) -0.04 (-0.11, 0.03)	**0.05 (0.00, 0.09)** 0.02 (-0.03, 0.06) 0.01 (-0.03, 0.06) -0.02 (-0.06, 0.03) 0.00 (-0.06, 0.05)	0.02 (-0.02, 0.06) 0.04 (-0.01, 0.09) 0.01 (-0.03, 0.05) -0.02 (-0.07, 0.02) 0.00 (-0.04, 0.04)	0.01 (-0.03, 0.04) 0.00 (-0.03, 0.03) -0.03 (-0.07, 0.00) 0.01 (-0.03, 0.05) 0.00 (-0.04, 0.03)
Political Orientation	-0.06 (-0.15, 0.02)	-0.04 (-0.12, 0.04)	-0.02 (-0.08, 0.05)	0.01 (-0.06, 0.08)	-0.01 (-0.08, 0.05)	0.01 (-0.03, 0.06)	0.00 (-0.05, 0.04)	0.00 (-0.03, 0.04)

	QRP 17	QRP 18	QRP 19	QRP 20	QRP 21	QRP 22	QRP 23	QRP 24
Item Wording	**Not maintaining comprehensive and detailed documentation of all decisions, modifications, and steps undertaken throughout the research process.**	**Not engaging in reflexivity, which involves failing to acknowledge or address my biases, perspectives, and potential influence on the research process.**	**Failing to sufficiently describe the methods of data collection and analysis, making it difficult for others to understand how conclusions were reached.**	**Phrasing questions in interviews or focus groups in a way that might influence participants towards certain responses.**	**Interpreting qualitative data in a way that confirms the analysts pre-existing ideas and beliefs, ignoring data that might suggest alternative interpretations.**	**Deliberately omitting or failing to discuss data that contradicts or does not align with the emerging themes or patterns identified in the research.**	**Presenting interview or focus group findings in a manner that may not fully reflect the original data.**	**Providing false or misleading information to participants about the purpose or nature of the study.**
Publication Pressure	0.07 (-0.02, 0.15)	0.04 (-0.04, 0.11)	**0.08 (0.01, 0.14)**	-0.02 (-0.09, 0.05)	0.01 (-0.04, 0.06)	0.02 (-0.03, 0.07)	0.06 (0.00, 0.13)	0.00 (-0.04, 0.04)
Funding Pressure	-0.06 (-0.17, 0.04)	**-0.08 (-0.15, -0.01)**	-0.03 (-0.09, 0.03)	0.01 (-0.06, 0.08)	0.03 (-0.02, 0.08)	-0.03 (-0.08, 0.02)	-0.05 (-0.11, 0.01)	-0.01 (-0.05, 0.03
Scientific Norms Universalism Communalism Organised Scepticism	-0.04 (-0.13, 0.04) 0.05 (-0.03, 0.12) -0.01 (-0.09, 0.07)	-0.04 (-0.11, 0.03) -0.01 (-0.07, 0.06) -0.07 (-0.15, 0.00)	**-0.12 (-0.17, -0.06)** -0.03 (-0.09, 0.03) **-0.10 (-0.17, -0.02)**	**-0.09 (-0.15, -0.03)** -0.02 (-0.08, 0.04) -0.07 (-0.15, 0.02)	**-0.05 (-0.10, -0.01)** 0.00 (-0.04, 0.04) **-0.11 (-0.17, -0.04)**	**-0.05 (-0.10, -0.01)** -0.03 (-0.07, 0.02) -0.07 (-0.14, 0.00)	**-0.09 (-0.14, -0.03)** -0.05 (-0.11, 0.00) -0.04 (-0.11, 0.04)	-0.02 (-0.06, 0.02) 0.00 (-0.03, 0.02) -0.02 (-0.07, 0.02)
Personal Motivations Career Growth Ego and Esteem	-0.05 (-0.14, 0.03) **0.13 (0.05, 0.22)**	-0.01 (-0.08, 0.06) 0.05 (-0.02, 0.12)	0.02 (0.02, 0.08) 0.02 (-0.04, 0.07)	-0.04 (-0.11, 0.03) 0.01 (-0.05,	-0.04 (-0.10, 0.01) **0.06 (0.01, 0.10)**	-0.01 (-0.06, 0.04) 0.01 (-0.04, 0.06)	0.00 (-0.07, 0.06) 0.05 (0.00,	0.00 (-0.03, 0.04) -0.01 (-0.05,

				0.08)			0.11)	0.02)
Proximal Science-Specific Motivations								
Impact and Quality	-0.06 (-0.15, 0.03)	-0.06 (-0.13, 0.01)	-0.03 (-0.10, 0.03)	0.00 (-0.06, 0.07)	-0.03 (-0.08, 0.01)	**-0.08 (-0.13, -0.03)**	**0.00 (-0.07, 0.06)**	0.00 (-0.04, 0.03)
Academic Achievement	-0.01 (-0.10, 0.08)	0.06 (-0.02, 0.13)	-0.04 (-0.10, 0.02)	0.04 (-0.03, 0.11)	0.01 (-0.04, 0.06)	0.03 (-0.01, 0.08)	0.03 (-0.03, 0.09)	-0.01 (-0.04, 0.01)
Research Approach								
Falsification	0.02 (-0.07, 0.12)	-0.05 (-0.13, 0.02	-0.05 (-0.11, 0.02)	**0.08 (0.01, 0.16)**	-0.02 (-0.07, 0.03)	**-0.06 (-0.11, -0.01)**	0.01 (-0.06, 0.08)	0.00 (-0.04, 0.03)
Confirmation	-0.09 (-0.19, 0.01)	-0.01 (-0.09, 0.07)	0.03 (-0.04, 0.09)	-0.02 (-0.09, 0.06)	0.03 (-0.03, 0.08)	**0.06 (0.01, 0.11)**	0.01 (-0.07, 0.09)	0.03 (-0.01, 0.07)
Personality Traits								
Extraversion	-0.01 (-0.07, 0.08)	-0.02 (-0.08, 0.05)	0.01 (-0.04, 0.07)	0.00 (-0.06, 0.07)	-0.02 (-0.06, 0.02)	-0.02 (-0.06, 0.03)	0.02 (-0.03, 0.07)	0.01 (-0.02, 0.05)
Agreeableness	0.02 (-0.05, 0.10)	0.02 (-0.05, 0.08)	-0.02 (-0.07, 0.03)	0.02 (-0.04, 0.07)	-0.01 (-0.05, 0.03)	-0.01 (-0.05, 0.03)	0.03 (-0.03, 0.08)	0.02 (-0.01, 0.06)
Conscientiousness	**-0.16 (-0.24, -0.07)**	**-0.11 (-0.19, -0.04)**	**-0.08 (-0.14, -0.01)**	-0.09 (-0.16, -0.01)	**-0.07 (-0.11, -0.02)**	-0.01 (-0.06, 0.03)	-0.04 (-0.10, 0.03)	**-0.06 (-0.10, -0.02)**
Neuroticism	-0.06 (-0.14, 0.02)	-0.02 (-0.08, 0.05)	0.02 (-0.03, 0.07)	**0.00 (-0.06, 0.06)**	-0.03 (-0.07, 0.02)	-0.02 (-0.06, 0.02)	-0.02 (-0.08, 0.04)	-0.02 (-0.05, 0.01)
Openness	**-0.11 (-0.20, -0.02)**	**-0.12 (-0.19, -0.05)**	-0.03 (-0.08, 0.03)	-0.05 (-0.13, 0.02)	**-0.05 (-0.10, 0.00)**	-0.02 (-0.07, 0.03)	-0.02 (-0.08, 0.04)	-0.01 (-0.05, 0.03)
Political Orientation	-0.06 (-0.14, 0.03)	-0.03 (-0.09, -0.09)	-0.06 (-0.11, 0.00)	-0.05 (-0.11, 0.01)	-0.02 (-0.07, 0.02)	-0.02 (-0.07, 0.03)	-0.01 (-0.08, 0.05	0.02 (-0.02, 0.06)

n = 3050; Coefficients and confidence intervals (CIs) are based on pooled estimates across 15 imputed datasets, using Rubin's rules. CIs are bootstrapped. Sampling weights were included in all models.
All models contain the four background characteristics of sex, methodological approach, career stage and country.
Bolded figures indicate statistical significance at the **0.05** alpha level.

Chapter 5

Discussion and Conclusion

Introduction

This thesis sought to explicate the factors and mechanisms that are associated with engagement in self-reported Questionable Research Practices (QRPs). The central research question guiding the empirical inquiry was: *In what ways are individual-, institutional-, and system-level factors associated with engagement in QRPs?* In addressing this question, the thesis explored how demographic, social, psychological, cultural, and systemic factors shape – and interact with – suboptimal researcher behaviour. It responded to an otherwise fragmented and disjointed body of literature, aiming to clarify and advance our understanding of *why* researchers compromise their integrity, cut corners and contribute to the canonisation of erroneous scientific findings.

Chapter 1 outlined the broader context for this inquiry, noting a growing consensus that the replication crisis and the lack of robustness in large segments of the scientific literature are less attributable to overt misconduct than to QRPs (though, the extent to which this relationship exists have been questioned; see, Ulrich & Miller, 2020; Rubin, 2023; Rubin, 2025; Francis, 2012).[9] Beyond issues of replication, QRPs can lead to the production of false research literatures, wasted resources, distorted theory development, poor educational signals and reduced credibility of legitimate findings. These practices are impactful not because of their severity in isolation, but because of their cumulative and systemic effects overtime. Yet, empirical investigation into why researchers engage in such practices – despite the idealised image of the scientist as objective, honest, and trustworthy – remains limited, fragmented, and underdeveloped. This thesis argues that further empirical inquiry into the correlates of QRPs is essential and that findings from such inquiry should be synthesised to support a more coherent understanding. A more robust evidence base is necessary for informing interventions,

[9] It is beyond the scope of this thesis to interrogate the downstream consequences of QRPs in any more detail, but I acknowledge that the degree to which QRPs are harmful to knowledge development is variable and not always clear.

policies, and governance structures that aim to mitigate engagement in QRPs. Such a foundation can support the correction of the scientific record, address the canonisation of erroneous findings, enhance the effectiveness of evidence-based solutions to societal problems, and foster public trust in science. Chapter 1 additionally introduced the methodological approach adopted to address these questions. The thesis employed both secondary analyses of large-scale secondary survey data (Papers A and B) and primary survey data collected as part of this research (Paper C).

In Paper A (Chapter 2), I assessed the extent to which self-reported QRP engagement varies across individual, organisational, and systemic levels using multi-level modelling on data from a large-*N* survey of researchers based in Europe and the United States. The survey used was collected as part of the Standard Operating Procedures for Research Integrity (SOPs4RI; Allum et al., 2023) project. The multi-level modelling approach enabled the partitioning of variance across individual, organisational, and national levels, thereby providing insight into the relative association of each level. It identified key individual-level predictors (e.g., career stage, contract type, disciplinary field, gender, and commitment to scientific norms) and organisational-level characteristics (e.g., research integrity policies, training, and culture).

In Paper B (Chapter 3), I used a Bayesian multi-level modelling approach to analyse a separate large-*N* dataset investigating whether perceptions of practice norms – specifically, the perceived normativity of QRPs – are related to individual self-reported engagement in QRPs, as well as the potential mechanisms of influence. Paper B used data from the Practices, Perceptions, and Patterns of Research Integrity (PRINT) project (further details contained in Schneider et al., 2024). Here, QRPs were treated as repeated measures nested within individuals, enabling estimation of the item-level relationships between perceived prevalence and self-reported engagement. It additionally explored the mechanisms through which prevalence estimates may influence self-reported QRP engagement. The first mechanism posits that perceiving a QRP to be descriptively normative, contributes to a belief that *not* compromising one's integrity will have a negative impact on one's career – that is, they will 'fall behind'. In turn, this motivates self-reported engagement in these practices. Furthermore, perceptions of peers' behaviours may exert a stronger influence on self-reported engagement in QRPs when those practices are perceived as particularly harmful to science. Conversely, if a specific QRP is not considered damaging, researchers might feasibly engage in it irrespective of their colleagues' behaviour.

Paper C (Chapter 4) extended this inquiry by constructing a comprehensive, theoretically informed model of self-reported QRP engagement, examining the interplay between motivations, traits, norms, and systemic pressures. It drew on a primary survey dataset of researchers who are operating across disciplinary fields to investigate correlates at both the individual and systemic levels. It explored whether proximal science-specific motivations mediated the relationship between personal motivations and self-reported QRP engagement and tested the associations with various systemic and psychological variables. Another key area of focus was whether Mertonian norms of science are differentially associated with QRP engagement, thus building on Paper B, where commitment to Mertonian norms were operationalised as a single, unitary construct. The process produced empirical results that cumulatively contribute to our understanding of the correlates of self-reported QRP engagement.

In this final chapter, I summarise the key empirical findings from across the three studies, outlining their implications for policy and governance, and consider how they can inform interventions aimed at improving research integrity. The limitations of the research are then discussed, before proceeding onto concluding remarks and a brief discussion of future inquiry in this area.

Summary of Empirical Findings

Paper A found that only a small proportion of the variance in self-reported QRP engagement can be attributed to differences between organisations, indicating that organisational-level factors explain relatively little variation in such behaviours, in comparison to individual-level differences. On this basis, we inferred that engagement in QRPs is feasibly attributable to individual-level and systemic-level factors. Papers B and C set out to explore these factors in more detail.

Systemic Factors: Publish or Perish

There is growing evidence that the scientific ecosystem as currently exists creates a criminogenic opportunity structure, inadvertently motivating engagement in QRPs. Paper B and Paper C showed that greater self-reported pressure to publish is associated with greater self-reported engagement in QRPs. This supports the dominant thesis in meta-science: that excessive pressure to publish compels researchers to cut corners to remain competitive and retain (or obtain) academic employment. This pressure is derived from assessments of researcher success being grounded in arbitrary metrics such as publication frequency, citation counts, and funding received. Such measures are not valid markers of quality (Fire and Guestrin, 2019). These reductive and misaligned measures of success stand in stark contrast to the core aims of science – namely, the advancement of certified knowledge (Merton, 1973, p. 270).[10] While a degree of incentivisation is not inherently problematic – and is, in fact, essential for shaping behaviour and aligning priorities in complex systems – the dominant incentive

[10] I avoid here getting into the philosophical discussions about what science *is* and what science *ought to be* but acknowledge that most delineations of the goals of science (Popper's, Merton's, Stokes, Pasteur's Quadrant) see 'knowledge accumulation' as at least one of its primary goals).

structures in science frequently fail to recognise or reward rigorous and responsible research practices. The rewards for engaging in QRPs can be substantial – ranging from securing permanent positions, tenure, and promotion, to gaining job security and increased, stable income. That is, the system *motivates* researchers to offend. The ethically ambiguous nature of QRPs, obscured by limited transparency – potentially by design – and reinforced by the relative autonomy of scientific practice, results in minimal oversight and a low probability of detection. This dynamic closely mirrors Merton's (1938) strain theory, wherein structural pressures to achieve culturally prescribed goals – such as publication, promotion, and prestige – can lead individuals to adopt illegitimate means when legitimate pathways are blocked or undervalued. In the current context, legitimate means – such as conducting rigorous, transparent research – are often devalued or obstructed by systemic pressures, while the cultural goals of productivity and prestige remain intact. As a result, some researchers adapt through 'innovation,' adopting illegitimate means – such as p-hacking or gift authorship – to achieve those goals. A growing body of evidence, including the empirical studies contained within this thesis, show self-reported publication pressure to be one of the strongest correlates of self-reported QRP engagement, relative to other explanatory factors. Specifically, in Paper A, I found that where researchers reported their working environments to be more collegial, without harmful publication pressure, detrimental power imbalances and conflict, self-reported QRP engagement was lower. Similarly, in Paper C, we found more direct evidence of this dynamic, with self-reported publication pressure being positively associated with self-reported QRP engagement. Strain theory, wherein structural pressures to achieve culturally prescribed goals – such as publication, promotion, and prestige – can lead individuals to adopt illegitimate means when legitimate pathways are blocked or undervalued.

Individual-level Factors

Although the 'publish or perish' thesis has empirical support, it presents an overly reductive account of poor research practice, potentially obscuring the role of individual agency and ethical responsibility. There is little room for doubt that structural pressures shape conduct, but they do not solely determine research behaviour. Scientists can exercise choice when navigating researcher degrees of freedom, even when incentives point them in a particular direction. This thesis has shown that researchers respond to pressures differently, depending on their values, motivations, normative commitments and personality traits. These individual-level factors are not only important, but they are also potentially more significant than structural components, such as unequal incentive structures. These individual-level factors will now be discussed in turn.

Mertonian Norms of Science

One of the most significant contributions of this thesis to the empirical literature pertains to the association of Mertonian norms with research practice. In Paper A, we find that greater commitment to the normative ideals of science is associated with decreased self-reported engagement in QRPs. This corroborates earlier, albeit limited, inquiry that finds an association between a commitment to scientific norms and engagement in QRPs (Gopalakrishna et al., 2022) and scientific misconduct (Anderson et al., 2007). In Paper C, we developed this by finding a distinct association of *each* Mertonian norm (namely, communalism, organised scepticism and universalism) and self-reported QRP engagement. In this study, 'scientific norms' were operationalised as distinct ideals, consistent with Merton’s original delineation of four normative principles. However, 'disinterestedness' – the fourth and final norm identified by Merton – was omitted from the analysis, as it did not emerge as a single, unified construct in the factor analysis. This suggests that disinterestedness may no longer represent a cohesive or widely endorsed scientific norm within contemporary research practice. A greater commitment to communalism, universalism and organised scepticism was found to be associated with reduced self-reported engagement in QRPs. Universalism was particularly interesting as it was a stronger predictor of self-reported QRP engagement than self-reported publication pressure. It was additionally found to moderate the relationship between publication pressure and self-reported QRP engagement, providing empirical evidence that a greater commitment to this ideal may buffer and safeguard against the influence of systemic pressures, such as publication pressure. These internalised standards help researchers navigate researcher degrees of freedom, act as buffers against perverse incentives and ensure that researchers maintain their epistemic integrity. It suggests that emphasis should be attributed to a researcher's value system, and that nurturing – or selecting researchers who are committed to – these values may serve to increase the robustness of the scientific enterprise.

Personal and Proximal Science-Specific Motivations

An additional key empirical contribution of this thesis is its identification that a researchers' personal and proximal science-specific motivations are associated with likelihood of engaging in QRPs – a dimension often overlooked in prior research. As presented in Paper C, we found that researchers' motivations are not only important determinants of self-reported QRP engagement, but in some cases, they are as equally important as systemic pressures. For example, where researchers are more motivated by ego and esteem (e.g., to achieve fame or avoid being seen as a failure), they are more likely to report engaging in QRPs. Moreover, where researchers are motivated to improve their career and ensure job security, they are more likely to engage in these practices. This provides empirical validation to parts of Jussim et al. (2018) social-psychological model of research practice, which posits that such personal motivations influence self-reported QRP engagement both directly and indirectly. Paper C found that the relationship with personal motivations operates, in part, through their association with proximal scientific motivations – namely, the pursuit of scientific success (e.g., increased publication output, higher citation counts) and impact- and quality-related goals (i.e., ensuring the scientific and societal relevance of one's work). While logical and perhaps commonsensical, this provides novel empirical evidence that a researcher's distinct motivations and values matter as much as systemic pressures in influencing researcher behaviour.

Personality Traits

The extant literature has demonstrated that personality traits are likely important correlates of self-reported QRP engagement (Schneider et al., 2024). Paper C found that researchers who are more agreeable, conscientious, emotionally stable and open report engage in QRPs less frequently. While effect sizes are small, it supports the concluding claims that a researcher's characteristics, including dispositional factors such as personality traits, are important in correlates of researcher conduct. Traits such as conscientiousness and agreeableness may buffer against systemic pressures, suggesting that fostering a culture of integrity may require attention not only to systemic reform, but also to supporting the development of individual ethical dispositions and norms. Again, this highlights the importance of integrating both individual-level and structural-level explanations when seeking to understand, predict, and ultimately mitigate self-reported engagement in QRPs.

Epistemic Approach

A further important, and novel, finding from Paper C concerns epistemic orientation. I found that where researchers seek to *falsify* findings and theories, the less likely they are to engage in QRPs. That is, setting out to test and potentially disconfirm theories and findings, in line with Popperian ideals of falsification, may act as a protective factor against QRPs. Conversely, I failed to find a relationship between the extent to which researchers seek to confirm their findings and self-reported QRP engagement. This asymmetry is noteworthy: while a falsification-oriented mindset appears to discourage self-reported engagement in QRPs, a confirmation-oriented mindset does not necessarily increase it. This may reflect the fact that seeking confirmation is widely accepted within the current incentive structure of science – particularly in contexts where novel, positive findings are rewarded – yet it does not inherently require the use of QRPs. However, an orientation toward falsification may impose higher internal standards for methodological rigour, transparency, and openness to disconfirming evidence, which collectively reduce the likelihood of engaging in QRPs.

Perceptions of Practice Norms

Perceptions of practice normativity may act as a means of neutralisation. In Paper B, I found that the more QRPs were perceived to be prevalent – or engaged in by colleagues – the greater the self-reported engagement in these practices. This supports findings elsewhere showing that the extent to which these practices are seen to be normal, the more ethically defensible they are perceived to be, and the more willing researchers are to engage in them (Sacco et al., 2018). It was hypothesised that the relationship between prevalence and engagement is mediated by a belief that not compromising integrity will influence one's career. However, we failed to support this hypothesis. Instead, I concluded that the influence of perceptions of normativity on self-reported QRP engagement potentially occurs through the process of 'informational social influence', whereby researcher's conform to the behaviours of others on the basis that the practices are judged to be appropriate and the individual has a desire to be correct (Festinger et al., 1950; Hardin & Higgins, 1996; Jones & Gerrard, 1967). Further enquiry is needed to establish the causal nature of this relationship and to unpack the mechanisms through which it operates. Nonetheless, these findings open the door to novel interventions – grounded in insights from social psychology and behavioural economics – that may be leveraged to shift research behaviour toward more robust and ethical practices (Hardin & Higgins, 1996; Jones & Gerrard, 1967).

Demography

There are clear demographic differences in self-reported QRP engagement. Across Papers A and C, we found that those early in their career report to engage in more QRPs compared to those in the middle, late or end of their careers. This is likely to do with the extent to which incentive structures, and in turn the pressures, are more likely to affect early-career researchers who are seeking jobs, permanency and tenure. That is, the motivation to cheat, and cut corners, is exacerbated by the precarity of their situation – they have more to gain, and more to lose. Similarly, across all three papers, it was found that those in the medical and natural sciences engage in more QRPs on average compared to those in the social sciences or humanities. This feasibly results from the comparably greater pressure that STEM researchers experience, compared to those in the social sciences, humanities and the arts. Certainly, bibliometric studies show systematic field differences in productivity, with greater productivity in medicine, and the natural and technology sciences – when 'whole counts' are being used, rather than when counts are fractionalised based on number of authors (Piro et al., 2013). Finally, we found that males are more likely to engage in QRPs than females. This is unsurprising, given that research shows that males have lower moral standards in competitive environments (Robinson et al., 2000) and are more likely to engage in deceptive practices (Friesen and Gangadharan, 2012).

Implications for Policy, Practice and Governance

Taken together, the empirical contributions presented across the three papers prompt critical reflection on how various stakeholders within the scientific ecosystem – including funders, publishers, and research institutions – should respond to the high prevalence of QRPs. These findings underscore the need for coordinated, multi-level responses that go beyond individual responsibility, addressing the structural conditions, incentive systems, and normative cultures that shape researcher behaviour. In this section, I describe the implications of the findings in this thesis for scientific policy, practice and governance.

Cumulatively, the findings demonstrate that both systemic pressures and individual-level motivations significantly shape the extent to which researchers engage in QRPs. The strong association with systemic pressures – particularly publication pressure – reinforces concerns that current academic incentive structures are misaligned with the goals of responsible and rigorous scientific practice. Evidence gathered through this dissertation, and mounting evidence from the extant literature, validates this relationship. Self-reported QRP engagement appears to be greater when publication pressure is greater. Interventions that aim to reduce QRP engagement must therefore address these underlying systemic conditions, including the overreliance on quantitative performance metrics in assessments for hiring, tenure, promotion and permanency. Instead, assessment metrics should account for quality of work, the extent to which scientific findings advance the canon of knowledge, and the real-world impact that studies have. Efforts are already taking place in this regard. Without reconfiguring these broader institutional incentives, meaningful change is unlikely to occur.

However, systemic reform on its own is unlikely to fully mitigate engagement in QRPs, as such practices are also shaped by individual motivations, epistemic orientations, and broader research cultures. Researchers motivated by career advancement, reputation, fame, and other ego-related goals appear more likely to compromise their research integrity. This tendency is, in part, driven by a pursuit of outcomes – such as increased citation counts and higher publication frequency – that are not aligned with the principles of robust and rigorous science. Similarly, particular personality traits, such as agreeableness and conscientiousness are associated with decreased self-reported engagement in these practices. This suggests that even in the absence of systemic pressures, internalised drivers may still lead to questionable research behaviour.

Importantly, these findings challenge the assumption that competency shortfalls and a misunderstanding of 'ethics' – particularly in its conventional, compliance-based form – is sufficient to reduce suboptimal research practice. This is supported by the finding in Paper B that self-reported engagement in research integrity training is statistically unrelated to extent of self-reported QRP engagement. Research elsewhere additionally fails to find a relationship between integrity training and QRP engagement (Anderson et al., 2007; Antes et al., 2010). The evidence suggests that many QRPs are not the result of ignorance or misunderstanding, but of strategic behaviour within a competitive system. Intentionality is important. As such, interventions should move beyond education alone and instead target the cultural and incentive-based roots of these motivations, including the symbolic and material rewards attached to visibility, productivity and prestige. The evidence suggests that many QRPs are not the result of ignorance or misunderstanding, but of strategic behaviour within a competitive system. Intentionality is important. As such, interventions should move beyond education alone and instead target the cultural and incentive-based roots of these motivations, including the symbolic and material rewards attached to visibility, productivity and prestige.

Another key finding is that researchers who report a stronger commitment to the normative ideals of science – such as, universalism, communalism and organised scepticism – are significantly less likely to engage in QRPs. This suggests that normative commitments may function as internal ethical compasses guiding decision-making and buffering against systemic pressures and self-interested motivations. Importantly, to at least some degree, these values may retain regulatory force even in environments where competitive pressures are high. On the one hand, this suggests that fostering a stronger cultural identification with the ethos of science may provide a protective effect. Rather than relying on formal ethics and integrity training, institutions may be better served by cultivating environments that affirm and reward commitments – such as through mentoring, recognition of transparent and open scientific practices, and leadership that models integrity. On the other hand, institutions could incorporate measures of a researcher's value system into recruitment, permanency and promotion assessment. If, as some research suggests, researchers acquire and develop their understanding of research integrity long before doctoral study (Satalkar & Shaw, 2019), efforts to instil these values only once individuals enter the workforce may be too late. On this basis, recruitment practices that prioritise alignment with integrity-related values may prove more effective than attempting to cultivate them post hoc. However, such an approach would raise clear concerns about academic freedom and the risk of privileging particular ideological or epistemic positions.

Finally, the limited role played by organisation-level and country-level factors suggests that institutional variation – while not irrelevant – is unlikely to be the primary driver of QRPs. This finding indicates that research integrity challenges are not the result of a few poorly governed institutions but are systemic in nature and broadly distributed. Local reforms, while useful, may be insufficient in isolation. Meaningful shifts in research culture will likely require coordinated efforts at both systemic and disciplinary levels. Learned societies, scholarly associations, and professional organisations are uniquely placed to initiate and sustain such change, given their role in shaping disciplinary norms, guiding ethical standards, and influencing the behaviour of their members. Collectively, these findings point to the need for multi-level, systemic responses that focus less on individual compliance, and more on cultivating values and reshaping structures that enable and reward engagement in QRPs.

Limitations

The three empirical papers contained within this dissertation advance our knowledge on the mechanisms and factors that are associated with self-reported engagement in QRPs. That being said, there are some distinct limitations that should be addressed. Foremost, the effect sizes across all papers, while consistent with the extant literature, are relatively small. On this basis, we should be concomitantly humble in our interpretations and not seek to overgeneralise or overstate the findings or conclusions drawn. Greater confidence in the validity of these findings will depend on replication and verification efforts, not only with similar populations but also across varied disciplinary, institutional, and cultural contexts. Certainly, the representativeness of the sample is affected by its composition – being drawn primarily from researchers in European and Anglophone countries – as well as by the use of a self-selected, non-probability sampling method in the study discussed in Paper C.

A further methodological limitation of this thesis lies in the cross-sectional design of the studies. While cross-sectional data allow for the identification of associations between variables at a given point in time, they do not permit causal inferences or capture how these relationships may evolve over time. For example, as stated in Paper B, it is possible that engagement in QRPs is antecedent to estimating the extent of QRP prevalence within one's field – that is, they may engage in QRPs, then extrapolate from their own behaviour when making judgements about the prevalence of that behaviour within their field. As such, longitudinal or experimental designs would be necessary to more robustly assess the directionality and temporal dynamics of the observed relationships across all studies in this thesis.

Self-report measures are subject to well-documented limitations, notably the tendency for respondents to under-report behaviours that are perceived as socially undesirable or ethically contentious. This may be especially salient in research on QRPs, where concerns about judgment or repercussions could influence disclosure. Research elsewhere has shown that estimates of QRP prevalence based on self-reported measures may be conservative (John et al., 2012). Across all studies, while anonymity in surveys was ensured, truth-telling was not incentivised, and therefore responses may be subject to bias. Further research within this domain should seek to mitigate social desirability and other biases by using novel methods of incentivisation.

The sensitive nature of the QRP items may have contributed to non-response bias. That is, researchers with a strong interest in research integrity and those who have been more willing to disclose/engage in QRPs may have been more likely to respond. Moreover, those who had previously engaged in QRPs, or perceive such disclosures as risky despite assurances of anonymity, may be less likely to respond. Therefore, across all studies, the sample may differ systematically from the population of interest. On this basis, prevalence estimates should not be perceived as precise estimates as they are feasibly influenced by differential association resulting from response bias. This non-response may have restricted variability in QRPs and predictors, and therefore results are indicative of associations within the sample, rather than population-level prevalence.

Relatedly, the small effect sizes are feasibly related to restrictions in range rather than the absence of substantively meaningful relationships. Likewise, small shifts in probability can have practically meaningful effects when aggregated across large numbers of researchers, studies and research institutions.

A final methodological limitation concerns the non-random selection of QRPs across the three surveys used in this thesis. The practices examined represent a specific subset that may qualitatively differ from those included in other studies. While care was taken to include QRPs relevant across methodological traditions and disciplinary contexts, the selected items do not constitute a comprehensive or fully representative inventory of QRPs. To partially address this issue, Paper C – which drew on primary data – tailored the list of QRPs to individual respondents based on their methodological orientation (i.e., qualitative, quantitative, or mixed-methods). This approach aimed to enhance the validity and relevance of the QRP measures by ensuring that respondents were only asked about practices applicable to their own research activities. In any case, the QRP items serve as a relatively blunt instrument for assessing a researcher's willingness to compromise their integrity.

While this thesis offers a comprehensive empirical analysis of the mechanisms and factors that motivate self-reported engagement in QRPs, it is not exhaustive. Several plausible causal mechanisms remain unexplored. For instance, the extent to which researcher competencies – such as methodological, statistical, or epistemological skillsets – influence QRP engagement was beyond the scope of this study. Although the findings highlight the importance of intention and motivation, it is also feasible that some researchers engage in QRPs not out of deliberate misconduct, but due to a lack of knowledge or understanding about how to avoid practices that compromise scientific rigour. Future research would benefit from investigating how deficits in training, supervision, or methodological literacy contribute to suboptimal research practice, and whether capacity-building efforts could serve as effective interventions.

Concluding Remarks and Future Directions

This thesis contributes to the explanatory framework of QRP engagement, identifying individual, organisational and systemic-level factors that feasibly influence engagement in self-reported QRPs. The findings suggest that systemic and individual-level factors are the most important correlates of self-reported QRP engagement, whereas organisational-level differences appear to play a comparatively minor role. A researcher's motivations, their commitment to the normative ideals of science as articulated by Robert Merton, and the pressures associated with publication collectively shape the incentive structures that may lead researchers to engage in practices that compromise the robustness, validity, and reliability of scientific work.

As noted, while this thesis offers a comprehensive analysis, it is by no means exhaustive. Several additional factors likely relevant to QRP engagement were not included in the explanatory models presented here. Further inquiry should seek to investigate the relative influence of other mechanisms – in doing so, more efficacious interventions can be developed to mitigate use of QRPs. Furthermore, verifying the causal direction of these relationships will require the use of experimental methods and longitudinal approaches capable of capturing change over time and isolating key causal mechanisms. We are only at the beginning of systematic inquiry into the complex and multifaceted drivers of QRPs, and researcher misbehaviour more broadly. While empirical progress has been made, only through sustained, interdisciplinary, and methodologically diverse inquiry can we develop a more complete and actionable understanding of the conditions under which research integrity is upheld or compromised. Through sustained inquiry, we can develop more effective strategies to strengthen scientific integrity – strategies whose implications extend well beyond the boundaries of academia, influencing the credibility of scientific knowledge, the efficacy of evidence-based policy, and public trust in science.

References

Allum, N., Reid, A., Bidoglia, M., Gaskell, G., Aubert-Bonn, N., Buljan, I., Fuglsang, S., Horbach, S., Kavouras, P., Marušić, A., Mejlgaard, N., Pizzolato, D., Roje, R., Tijdink, J., & Veltri, G. (2023). Researchers on research integrity: a survey of European and American researchers. *F1000Research*, *12*, 187. https://doi.org/10.12688/f1000research.128733.1

Anderson, M. S., Horn, A. S., Risbey, K. R., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a National Survey of NIH-funded scientists. *Academic medicine : journal of the Association of American Medical Colleges*, *82*(9), 853–860. https://doi.org/10.1097/ACM.0b013e31812f764c

Anderson, M. S., Martinson, B. C., & De Vries, R. (2007). Normative dissonance in science: results from a national survey of u.s. Scientists. *Journal of empirical research on human research ethics : JERHRE*, *2*(4), 3–14. https://doi.org/10.1525/jer.2007.2.4.3

Antes, A. L., Wang, X., Mumford, M. D., Brown, R. P., Connelly, S., & Devenport, L. D. (2010). Evaluating the effects that existing instruction on responsible conduct of research has on ethical decision making. *Academic medicine : journal of the Association of American Medical Colleges*, *85*(3), 519–526. https://doi.org/10.1097/ACM.0b013e3181cd1cc5

Festinger, L. (1950). Informal social communication. *Psychological Review, 57*(5), 271–282. https://doi.org/10.1037/h0056932

Fire, M., & Guestrin, C. (2019). Over-optimization of academic publishing metrics: observing Goodhart's Law in action. *GigaScience*, *8*(6), giz053. https://doi.org/10.1093/gigascience/giz053

Friesen, L., & Gangadharan, L. (2012). Individual level evidence of dishonesty and the gender effect. *Economics Letters*, *27*(3), 624-626. https://doi.org/10.1016/j.econlet.2012.08.005

Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PloS one*, *17*(2), e0263023. https://doi.org/10.1371/journal.pone.0263023

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition, Vol. 3. The interpersonal context* (pp. 28–84). The Guilford Press.

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. https://doi.org/10.1177/0956797611430953

Jones, E.E., & Gerard, H. (1967). *Foundations of social psychology.* Wiley.

Jussim, L., Krosnick, J. A., Stevens, S. T., & Anglin, S. M. (2019). A Social Psychological Model of Scientific Practices: Explaining Research Practices and Outlining the Potential for Successful Reforms. *Psychologica Belgica*, *59*(1), 353–372. https://doi.org/10.5334/pb.496

Merton, R. K. (1938). Social Structure and Anomie. *American Sociological Review*, *3*(5), 672-682. https://doi.org/10.2307/2084686

Merton, R. K. (1973). The Normative Structure of Science. In Storer, N. W. (Ed.), *The Sociology of Science Theoretical and Empirical Investigation* (pp. 223-267). The University of Chicago Press.

Piro, F. N., Aksnes, D. W., & Rørstad, K. (2013). A macro analysis of productivity differences across fields: Challenges in the measurement of scientific publishing. *Journal of the American Society for Information Science and Technology*, *64*(2), 307-320. https://doi.org/10.1002/asi.22746

Robinson, R. J., Lewicki, R. J., & Donahue, E. M. (2000). Extending and Testing a Five Factor Model of Ethical and Unethical Bargaining Tactics: Introducing the SINS Scale. *Journal of Organizational Behavior*, *21*(6), 649–664. http://www.jstor.org/stable/3100394

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In Defense of the Questionable: Defining the Basis of Research Scientists' Engagement in Questionable Research Practices. *Journal of empirical research on human research ethics : JERHRE*, *13*(1), 101–110. https://doi.org/10.1177/1556264617743834

Satalkar, P., & Shaw, D. (2019). How do researchers acquire and develop notions of research integrity? A qualitative study among biomedical researchers in Switzerland. *BMC medical ethics*, *20*(1), 72. https://doi.org/10.1186/s12910-019-0410-x

Schneider, J. W., Allum, N., Andersen, J. P., Petersen, M. B., Madsen, E. B., Mejlgaard, N., & Zachariae, R. (2024). Is something rotten in the state of Denmark? Cross-national evidence for widespread involvement but not systematic use of questionable research practices across all fields of research. *PloS one*, *19*(8), e0304342. https://doi.org/10.1371/journal.pone.0304342