

Article

Reflexive Behaviour: How Publication Pressure Affects Research Quality in Astronomy

Julia Heuritsch

Research Group “Reflexive Metrics”, Institut für Sozialwissenschaften, Humboldt Universität zu Berlin, Unter den Linden 6, 10117 Berlin, Germany; julia.heuritsch@hu-berlin.de

Abstract: Reflexive metrics is a branch of science studies that explores how the demand for accountability and performance measurement in science has shaped the research culture in recent decades. Hypercompetition and publication pressure are part of this neoliberal culture. How do scientists respond to these pressures? Studies on research integrity and organisational culture suggest that people who feel treated unfairly by their institution are more likely to engage in deviant behaviour, such as scientific misconduct. By building up on reflexive metrics, combined with studies on the influence of organisational culture on research integrity, this study reflects on the research behaviour of astronomers with the following questions: (1) To what extent is research (mis-)behaviour reflexive, i.e., dependent on perceptions of publication pressure and distributive and organisational justice? (2) What impact does scientific misconduct have on research quality? In order to perform this reflection, we conducted a comprehensive survey of academic and non-academic astronomers worldwide and received 3509 responses. We found that publication pressure explains 10% of the variance in occurrence of misconduct and between 7% and 13% of the variance of the perception of distributive and organisational justice as well as overcommitment to work. Our results on the perceived impact of scientific misconduct on research quality show that the epistemic harm of questionable research practices should not be underestimated. This suggests there is a need for a policy change. In particular, lesser attention to metrics (such as publication rate) in the allocation of grants, telescope time and institutional rewards would foster better scientific conduct and, hence, research quality.

Keywords: reflexive metrics; publication pressure; research quality; research culture; research integrity; research misconduct



Citation: Heuritsch, J. Reflexive Behaviour: How Publication Pressure Affects Research Quality in Astronomy. *Publications* **2021**, *9*, 52. <https://doi.org/10.3390/publications9040052>

Academic Editor: Bart Penders

Received: 10 September 2021

Accepted: 3 November 2021

Published: 9 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The growing body of research on the effect of evaluation procedures on scientific behaviour (e.g., [1–4]) points towards performance indicators (such as publication and citation rates) not only describing, but also prescribing behaviour [5,6]. In other words, they have *constitutive effects* on the knowledge production process [7]. This suggests that metrics intended to measure concepts like research quality, end up *defining* what research quality means, and thereby *shaping* what researchers strive for. As Dahler-Larsen states: “A claim to measure quality cannot be understood as referring to an already-existing reality, but as an attempt to define reality in a particular way” [8] (p. 11). Metrics are therefore not merely proxies for quality, but also represent a definition of how quality is considered.

Capturing a complex concept, such as research quality, quantitatively strips it of its complexity. This makes it easier to understand, turning it into something objective and comparable [5,8], but also leads to a “validity problem” [7] (p. 971). The indicator’s inability to account for the phenomenon’s full complexity may therefore lead to an “evaluation gap” [9]. This may lead to unintended consequences when putting indicators into place, such as *scientific misconduct* resulting from *coping* with the divergence between what quantitative proxies measure and what researchers value themselves [10].

Which metrics are considered important depends on the *culture of science*, which has evolved through neoliberal reforms and the rise of new public management in the last

30 years [11]. Researchers have become increasingly dependent on external resources, such as funding and rewards. *Competition* for these resources and for positions has intensified [12]. *Publish or perish* is an integral part of this culture, since a scientist's reputation (arguably the most important currency in academia), their funding opportunities and their career development hinge on their metrics, such as their publication rate [13,14]. As a result, researchers have an interest in scoring well on performance indicators. Due to this "goal displacement", where doing well on quantitative metrics becomes an aim in itself [4] (p. 27), researchers may adopt various gaming strategies to attain the goal (refer to, e.g., [3,15]). This suggests that an indicator ceases to be a good measurement when it becomes a target (Goodheart's Law).

Adopting gaming strategies to hit a target set by performance indicators is what Fochler and De Rijcke [4] call playing the "indicator game". While some forms of gaming may seem innocent at first (e.g., going for an easy publication), they can result in behaviour that scientists themselves perceive as a threat to *research integrity*. This may range from questionable research practices (QRPs), such as insufficient supervision of (graduate) students or salami slicing to publish more papers on one's research, to outright scientific misconduct such as fabrication, falsification or plagiarism (FFP; [16]). Martinson et al. [17] have shown that the latter, more extreme kinds of misconduct, are less frequent than the "'minor offences', the many cases of 'sloppy science'" [18] and "carelessness" [19] (p. 2), represented by the QRPs. Since they are more numerous and more difficult to spot, the authors of [17] (p. 737) suggested that such "mundane 'regular'" forms of misbehaviour pose larger threats to research integrity than outright fraud. If playing the indicator game is at least partly causing research misconduct, it follows that using indicators in research evaluation has an impact on *research integrity*.

Studies on the constitutive effects and unintended consequences of indicator use on research (behaviour) are designated under the umbrella term "*reflexive metrics*" [20] (p. 146). The *relationship between research integrity and research culture and climate* has also been studied in literature (e.g., [21–24]). Academic culture may comprise of networks of peers, departments, institutions, funding agencies, grant reviewers, journal editors and the peer-review system [19]. These authors carried out the first systematic, quantitative analysis of the relationship between organisational culture, perceptions of justice and scientist's behaviours. They found evidence for greater perception of injustice leading to misbehaviour, especially among researchers whose career is at stake (e.g., early-career researchers). Other studies related individual perceptions of research climate, such as advisor–advisee relations or expectations, to misconduct (e.g., [21]). Anderson et al. [12] and Martinson et al. [25] found evidence that the greater competition resulting from the neoliberalist culture in science resulted in gaming strategies to the detriment of research integrity.

Research integrity has been linked to *research quality* (e.g., [21,26]), and the former is easier to measure than the latter [23]. As both have a direct correlation with scientific misconduct, and a climate of research integrity fosters science quality, both terms are often equated in the literature on the effects of cultural aspects on research integrity. Crain et al. [21] (p. 837) point out that improving research quality is in fact the "holy grail" of organisational initiatives targeted at an ethical organisational climate. Given that metrics have constitutive effects, simply "fixing" aspects of quality by an indicator [8] (p. 143) will likely not lead to this holy grail. Instead, quality indicators "are neoliberal instruments which colonise practices and undermine professional values" [8] (p. 14). Since directly measuring research quality is therefore unfeasible, studying cultural aspects about a research environment seems to be the way forward to foster scientific quality.

The effects of *publication pressure* on research quality have received particular attention in recent studies, since publication rate is one of the key metrics in academia (cf. [2,14]). Haven et al. [27] summarise some key studies' findings: while some extent of publication pressure may be a driver for productivity, too much of it may not only have negative effects on research integrity and quality, but also on individual researchers. Examples of these negative effects include secrecy (e.g., a lower willingness to share data; [28]), less academic

creativity, less reliable science, neglect of negative findings [10,20] and a greater likelihood to engage in misbehaviour (QRPs and FFPs; [29–31]). The perceived competition resulting from the *publish-or-perish* imperative may also lead to emotional exhaustion on the individual level and feelings of unworthiness [20,32,33]. However, previous studies on the effects of publication pressure on research quality have two shortcomings: (1) Quantitative studies have thus far mainly included scientists from specific disciplines, such as biomedicine, management and population studies, and of specific academic ranks [31,34,35]. (2) While previous literature acknowledges the link between research integrity and research quality, we are currently lacking quantitative studies that explore the impact of misbehaviour on research quality as opposed to the impact on research integrity. Previous studies merely imply the impact of scientific misconduct on research quality, caused by compromised research integrity. (3) Haven et al. [36] conducted the only quantitative study to our knowledge that studies both the impact of publication pressure and the impact of research climate on research integrity.

To address these shortcomings, this paper aims to study quantitatively the impact of research culture on research quality in a natural science field—astronomy. The aspects of research culture under scrutiny are as follows: perceived publication pressure, perceptions of distributive justice and procedural justice in peer review, grant application and telescope time application processes. These cultural aspects have been found relevant in astronomy in a qualitative study [10]. Heuritsch [10] studied the “organisational hinterland” [8] of astronomy to understand how quality inscriptions are produced, how they diverge from the astronomers’ definition of quality and how this discrepancy affects research behaviour. In a nutshell, Heuritsch [10] found evidence for the structural conditions in the field, especially the overemphasis on performance measured by publication rate, reception of external grants and telescope time, leading to gaming strategies to score well on those indicators mostly in the form of QRPs. These are found to be a response to the dissonance between cultural values (producing qualitative research that genuinely pushes knowledge forward) and the institutional objectives imposed to have career in academia (scoring well on indicators). In other words, there is a discrepancy between what indicators measure and the astronomers’ definitions of scientific quality—the so-called evaluation gap. Gaming strategies then give the appearance of compliance with cultural values, while using institutionalised means to achieve a good bibliometric record in innovative ways, such as salami slicing, cutting corners or going for easy publications [37]. Haven et al. [37] found evidence for a decrease in overall research quality as a consequence of prioritising quantity.

Based on Heuritsch [10,20], we can use astronomers’ own definitions of research quality, as well as previous studies on the relationship between academic culture and research behaviour, to analyse the effect of perceived publication pressure and organisational justice on research behaviour **and** quality in astronomy. Understanding which cultural aspects foster and which inhibit research quality in this field will bring us a step closer towards the holy grail—the knowledge of how to support the scientific enterprise.

This is not only the first study to analyse the effects of cultural aspects (such as publication pressure) on research quality in astronomy, but also the first to integrate reflexive metrics and studies on the relationship between research culture and integrity. Moreover, it is the first to employ structural equation modelling to fully account for the structural relationships between the phenomena of interest.

This paper is structured as follows. First, we give a theoretical background on explanations of misconduct. The aim of that section is to forge a bridge between reflexive metrics and theories of scientific misconduct, many of which take organisational culture into account. We demonstrate how rational choice theory integrates these partial theories of misconduct and therefore provides a suitable theoretical framework that sets the basis for the constructs we used in our statistical analysis and the relationships we test among them. Second, the Methods Section describes the sample selection, the survey instruments, research question and hypotheses and the technical aspects with regards to how we performed the statistical analyses. Third, the Results Section contains descriptive statistics,

the results from our EFAs, CFAs and SEM with scientific misconduct as the dependent variable, as well as the perceived impact of scientific misconduct on research quality. The Results Section is followed by a Discussion Section, Strengths and Limitations Section and a Conclusions Section that also gives an outlook for future studies.

2. Theoretical Background: Explanations for Misconduct

To understand metrics' role in research misconduct (hereafter referred to as *misconduct* or *misbehaviour*), we must first reflect on potential causes of misbehaviour. Haven & van Woudenberg ([18] (based on [38])), suggest that there are three 'narratives' that may help us understand misconduct:

- (i) Failures on the individual actor's level ("impure individuals");
- (ii) Failures on an institutional level (of a particular university/institute);
- (iii) Failures on the structural system of science level.

Haven and van Woudenberg [18] point out that these narratives are not mutually exclusive, and test six theories taken from previous literature on misconduct to assess their value in explaining one or more of these narratives. Five of those six theories shall be considered for this study:

- (1) Bad apple theories;
- (2) General strain theory;
- (3) Organisational culture theories;
- (4) New public management; and
- (5) Rational choice theory.

While a comprehensive review of these theories is outside the scope of this paper, more background information on this discourse can be found in Martinson et al. [19] and Haven and van Woudenberg [18]. To set the theoretical background for this story, we shall give a brief overview and describe how (1) to (4) can be subsumed under (5)—rational choice theory.

Bad apple theories provide perhaps the earliest explanations for misconduct [39] and account for the first narrative (i), since misbehaviour is thought to be solely caused by an individual and their distorted psychology. However, these theories are regarded as too simplistic in a sociological context, as they do not account for any institutional (ii) or structural (iii) contexts [22].

General strain theory (GST; [40]) is an "important strand of deviance theory, as it is the pressure to deviate from accepted norms as a response to perceived injustice" [19] (p. 4). Based on Durkheim's concept of anomie, Merton [41] suggests that deviant behaviour may be a *coping response* to structural strain, resulting from the inability to meet cultural ends with culturally legitimate means. The deviant behaviour (such as scientific misconduct) motivated by this kind of stress is therefore nothing else but an "innovative" pathway to success. Agnew's GST enhances Merton's concept with the idea that deviant behaviour is not a necessary outcome of strain, but that coping strategies also depend on individual traits, such as self-esteem and intelligence, and contextual factors, such as a strong social support network and whether peers show legitimate or illegitimate coping behaviour. Since GST recognises individual–environment interactions and strain resulting from structural conditions, it accounts for all three narratives.

Organisational culture theories (OCTs), rooted in organisational psychology, recognise that the culture and structure of the organisation in which an individual works affects their behaviour. In the context of academia, the organisational structure spans local (e.g., departments or research institutes) and external settings (e.g., funding agencies, [inter-]national peer-review systems, the overall academic employment market, etc.; [19]). Organisational culture encompasses all explicit and implicit norms and values within the organisation. A particular strand of OCT, organisational justice theory (OJT), suggests that individuals who perceive they are being treated fairly by their organisation, behave more fairly themselves [19,26] (These authors also give more information on the development of

OCTs and OJTs, which would go beyond the scope of this paper [19]). The fairer people feel their organisation's processes are, the more likely they are to trust their workplace, to comply with decisions made and to not engage in questionable behaviour (19,26). In other words, people who perceive the distribution of resources and decision-making processes as fair are more likely to respond with normative as opposed to deviant behaviour, such as scientific misconduct [19]. One may distinguish between two types of organisational justice (OJ; *ibid.*): procedural and distributive justice. The former refers to a perception of fairness in decision-making processes and the latter to fairness in resource distribution processes. In academia, these processes may stem from the local and external settings of the organisational structure, such as the peer review of manuscripts, tenure, promotion and peer-review committees for research grant proposals [26]. Since OCTs recognise the fact that characteristics of the environment in which researchers work promote or inhibit scientific integrity, they are institutional (ii) and structural (iii) narratives.

New public management (NPM) is a form of public administration based on neoliberal policies and is characterised by a combination of free market ideology and intense managerial control practices [11]. The author (p. 601) poses the formula: "free market = competition = best value for money = optimum efficiency [. . .]". Arguably, this formula is the rationale for competition in science, since NPM practices have reached academia since the 1980s (*ibid.*). The NPM paradigm also values efficiency as a key objective, as evident from the formula. The call for increased accountability in science (cf. [42]) can be associated with this strive for efficiency. Accountability is in turn sought using quantitative performance indicators, such as publication rates and impact factors. Since resources are limited, there are less tenured positions in science than there are graduate students, and given the extreme focus on efficiency and performance, NPM may result in "hypercompetition" in academia [43] at the cost of research integrity and the scientific enterprise (Anderson et al., 2007).

"Academic misconduct is considered to be the logical behavioral consequence of output-oriented management practices, based on performance incentives." [44] (p. 1140).

Therefore, the NPM theory of scientific misbehaviour suggests that if there is an over-emphasis on performance and competition, researchers will tend towards "self-protective and self-promoting" behaviours such as mistrust in peers, an aversion towards sharing information and data, QRPs and FFPs [12] (p. 459). This theory falls under the structural system of science narrative (iii).

Rational choice theory (RCT) is proposed by Heuritsch [10] as a suitable framework to study the emergence and the impact of the evaluation gap in the academic field of astronomy. Despite RCT's roots in economics, and contrary to what is often stated in the literature (e.g., [36,45]), when applied thoroughly in sociology (according to Esser [46]), RCT does not suggest that individuals act "rationally" in the classical sense (i.e., high investment of cognitive resources and disregarding emotions/instincts) or the simplistic "Homo Economicus". Instead, RCT explains how it is often more rational not to invest cognitive resources and instead follow scripts, which are predefined instructions on how to act in situations based of cultural and/or institutional norms [46]. The thorough application of RCT follows the Coleman Boat [47] by first analysing the *logic of an actor's situation*, which is comprised of one internal and three external constituents: (1) the internal component, made up of the actor's values, drivers, skills and personality; (2) the material opportunities at present; (3) explicit and implicit institutional norms; and (4) the cultural reference frame, such as symbols and shared values. This *situation* is in a second step translated into *bridge hypotheses*, which deliver the variables that are relevant for the individual's action that follows from the *situation*. An *action theory* (oftentimes the expected utility theory; (cf. [10,46])) is applied to explain how choices are made based on the derived variables of the *situation*. Third, by finding suitable *transformation rules*, one can explain how individual actions aggregate to the sociological phenomenon in question. The result is a sociological explanation of how the interplay between structural conditions, institutional norms and

the individual's personalities cause collective social phenomena. Applied to explaining scientific misbehaviour, RCT therefore pays tribute to all three narratives.

Building up on Heuritsch [10], which followed the Coleman Boat in their analysis, we find that theories (1) to (4) may be subsumed under RCT. First, the internal constituent of the individual's *situation* accounts for the person's beliefs, values and drivers. If they do not correspond to the cultural values, such an individual may assume the role of a "bad apple" (1) in the respective system/organisation. Second, if the cultural values cannot be lived up by institutionally legitimate means, the individual may perceive strain in the form of anomie (2; cf. [10]). Third, the external constituents of the *situation* account for the organisational culture and environment (3), and therefore its subsequent relevance in the individual's action. For example, perceived injustices may contribute to strain (2; [26]). Fourth, the prevailing NPM paradigm (5) influences the organisational culture (3), and therefore explains part of the culture's values and norms [36]. For example, NPM is likely to foster hypercompetition, which may impose strain (2), which in turn may lead to deviant behaviour, depending on individual's dispositions (1) that regulate the response to strain. This is how theories (1) to (4) deliver partial explanations for misbehaviour, which subsumed under RCT (5) can achieve a wider explanatory value.

As outlined in the introduction, NPM's focus on efficiency and accountability makes performance metrics an important aspect of the culture in science. Resource distribution and career choices made on basis of quantitative indicators may shape how fair researchers perceive the organisational culture, which in turn shapes research behaviour. Further, an overemphasis on metrics in research evaluation may put pressure on individuals to perform well on those indicators, which may result into coping responses to this strain. Research behaviour, thus, is reflexive insofar, that it is shaped by the culture actors are situated in. Reflexive metrics focusses on how behaviour depends on the way one's performance is evaluated. The constructs measured in this survey and the relationships we test among them are derived from this theoretical background and previous research based on the theories therein.

3. Methods

3.1. Sample Selection and Procedure

Astronomy is a highly international and globalised field [10,48]. This includes a high proportion of international collaborative research [49], not least because of the sharing of observatories located in specific parts of the world (e.g., ALMA in Chile). Astronomers have a strong common culture through their publication systems (three main journals and conference proceedings), societies and professional associations [10,48]. Heidler [50] estimates there to be about 15,000–20,000 active astronomers worldwide. They may work in academic research or in other research facilities, such as space agencies or non-public institutes. Given the close similarity in culture of academic and non-academic astronomers, this survey targeted both groups.

The ideal aim was to run a census. However, there is no official, complete list of all astronomers worldwide. Therefore, we used a three-stage cluster sampling technique to build our sampling frame, encompassing as many astronomers as possible. In the first stage, we constructed a list of astronomy institutions worldwide, including universities, non-academic research organisations, observatories, societies and associations. In the second stage, based on this list, we reached out to 176 universities, 56 non-academic research facilities and observatories and 17 societies and associations. We did so by emailing a respective contact person (e.g., a secretary/department head), including an invitation to the survey and asking them to forward the email to all department members (including PhD students). We estimate that 1200 academic astronomers were reached in this way. In the third stage, we contacted the heads of the nine divisions (https://www.iau.org/science/scientific_bodies/divisions; accessed on 10 September 2021) of the International Astronomical Union—the largest association of academic and non-academic astronomers—asking them to forward the invitation to their respective division members. Three of them followed

our request and one posted the invitation in their newsletter. In the fourth and final stage, we used an automated script to send out email invitations to the five remaining divisions, whose heads were non-responsive, using publicly available email addressees. That way, all IAU members (<https://www.iau.org/administration/membership/member/>; accessed on 10 September 2021) (approximately 12,000 astronomers) were reached through at least one channel. Note, that some astronomers may have received the invitation more than once, since astronomers may be part of more than one division and may also have been reached through multiple of the approaches described above. Given that the IAU only has around 7% of “junior members”, we asked survey recipients to forward the invitation to early-career researchers. We estimate that around 13,000–15,000 astronomers were reached in total, 3509 of which at least partly completed the survey, amounting to a response rate of roughly 25%. In total, 2011 astronomers completed the survey in full.

3.2. Instruments

We used the online tool LimeSurvey to create and host the survey. As outlined above, the survey is embedded in the conceptual framework of rational choice theory (RCT), subsuming organisational culture theory (OCT) and its subtype organisational justice theory (OJT). Therefore, our online survey contains a number of instruments to measure our independent and dependent variables:

3.2.1. Dependent Variables

- Scientific (Mis-)Behaviour

While there is no general definition of research misconduct, since it is highly contextual [1], Martinson et al. [15,19,25,26] and Bouter et al. [51] designed questionnaires asking about the occurrence of, in total, 90 different misbehaviour and questionable research practices. Based on Heuritsch [10], we chose the 18 items that were most relevant for the context of astronomy for our study (see Table S1a,b in S1). In our survey, we ask about the perceived frequency of observed (mis-)behaviour.

- Research Quality

We operationalised research quality in astronomy on the basis of the findings of Heuritsch [10,20]. The author found three quality criteria: (1) good research needs to push knowledge forward, which includes studying a diversity of topics and making incremental contributions; (2) the research needs to be based on clear, verifiable and sound methodology; and (3) the research needs to be reported in an understandable and transparent way. This includes the sharing of data and reduction codes. In the line of Bouter et al. [51], who surveyed the frequency of misbehaviour and its impact, for each (mis-)behaviour item we asked about the frequency (as mentioned above), the impact on the validity of the findings and the impact on the communication value of the resulting paper, and for two items we additionally asked for the impact on the research diversity (see Table S1a,b in S1). In analogy to Bouter et al. [51] (p. 2), “the total harm [on quality] caused by a specific research misbehaviour depends on the frequency of its occurrence and the impact [on quality] when it occurs”.

3.2.2. Independent Variables

- Perceived Publication Pressure

To measure perceived publication pressure, which has been linked with (perceived) misbehaviour [36], we adapted the Publication Pressure Questionnaire (PPQ; as validated by Tjldink et al. [13]) to the context of research in astronomy. The initial PPQ consists of 18 items, and we added 4 more. The added questions dealt with the influence of perceived publication pressure on the publication of data, reduction algorithms and replicability—all three of which have been found important for research quality in astronomy [10,20]. The 22 adapted PPQ items can be found in Table S2 in S1.

- Perceived Organisational Justice: Distributive and Procedural Justice

We measured perceived distributive justice via the Effort–Reward Imbalance (ERI; [52]) instrument. We adapted the short version, which consists of 10 items, and added 1 more item, resulting in 3 effort and 8 reward items (see Table S3 in S1). With regards to perceived procedural justice, in this study, we considered the following processes: (a) resource allocation, (b) peer review, (c) grant application and (d) telescope time application. We adapted the instruments Martinson et al. [19] used to the context of astronomy and the specific processes (see Table S4a,d in S1). In particular, we added questions about how much the success of that process depends on luck, on the one hand, and on improper preferential treatment, on the other hand. The addition of these two items followed from findings by Heuritsch [10] and suggestions from initial tests with astronomers.

- Perceived Overcommitment

Perceived overcommitment (OC), or the inclination to overwork, may be positively related to misbehaviour [19]. We adapted the six overcommitment items contained in the ERI instrument [52] (see Table S5 in S1).

3.2.3. Control Variables

In addition to our main independent variables that may predict misbehaviour, we assume that role-associated and individual aspects factor into these relationships. On the basis of Heuritsch [10], we expect that scientists who have not yet established their reputation, such as early-career researchers, will be more likely to perceive publication pressure and organisational injustice, as there are insufficient tenured positions. As opposed to Martinson et al. [19], who suggest that role-associated aspects (“social identity”) mediate the relationship between the independent and the dependent variables, we propose that they predict the independent variables. Our control variables include the following: gender, academic position, whether one is primarily employed at an academic (as opposed to non-academic) institution, whether one is employed at an institution in the Global North/South (https://meta.wikimedia.org/wiki/List_of_countries_by_regional_classification; accessed on 10 September 2021) and the number of published first or co-author papers in the last 5 years. The reference categories are as follows: gender: female/non-binary; academic position: full professor; institute location: Global South; number of papers published: 1–5.

3.3. Research Question and Hypotheses

In light of the theoretical background as outlined above, we work from the assumption that higher perceived procedural and distributive injustice in research and higher perceived publication pressure will increase scientists’ chances to observe research misbehaviour. We expect early-career researchers, with a less secure position, to be more likely to perceive both injustice and publication pressure.

Our research question is as follows: To what extent can role-associated factors, cultural aspects and publication pressure explain the variance in perceived research misbehaviour, and what effect does misbehaviour have on the research quality in astronomy?

Building on the qualitative study of the evaluation gap and its potential consequences on research quality in astronomy by Heuritsch [10] and previous studies on the relationship between research culture and integrity (e.g., [19]), this study tests the following hypotheses by means of a quantitative survey:

1. (H1): The greater the perceived *distributive injustice* in astronomy, the greater the likelihood of a scientist observing misbehaviour.
2. (H2): The greater the perceived *organisational injustice* in astronomy, the greater the likelihood of a scientist observing misbehaviour.
3. (H3): The greater the perceived *publication pressure* in astronomy, the greater the likelihood of a scientist observing misbehaviour.
4. (H4): Those for whom injustice and publication demands pose a more serious threat to their academic career (e.g., early-career and female researchers in a male-dominated field) will perceive the organisational culture to be more unjust, the publication pressure to be higher, and subsequently there will be more occurring misbehaviour.

5. (H5): Scientific misbehaviour has a negative effect on the research quality in astronomy.
6. (H6): The greater the perceived *publication pressure* in astronomy, the greater the likelihood of a scientist perceiving a greater distributive and organisational injustice.

Based on these hypotheses, we specified our structural equation model. Given that publications are the main output of research and one of the key indicators used to evaluate the performance of astronomers [10], we hypothesise that the perception of distributive and organisational justice depends on the perceived publication pressure. Overcommitment to work may depend on the perceived publication pressure as well as perceived distributional and organisational justice. We test the influence of all our control variables (academic position, gender, number of published papers, being academic/non-academic, employment in the Global North/South) on all our independent variables and the dependent variable, scientific misconduct. Figure 1 depicts the level of the latent constructs of this model, excluding the measured indicators, for simpler readability.

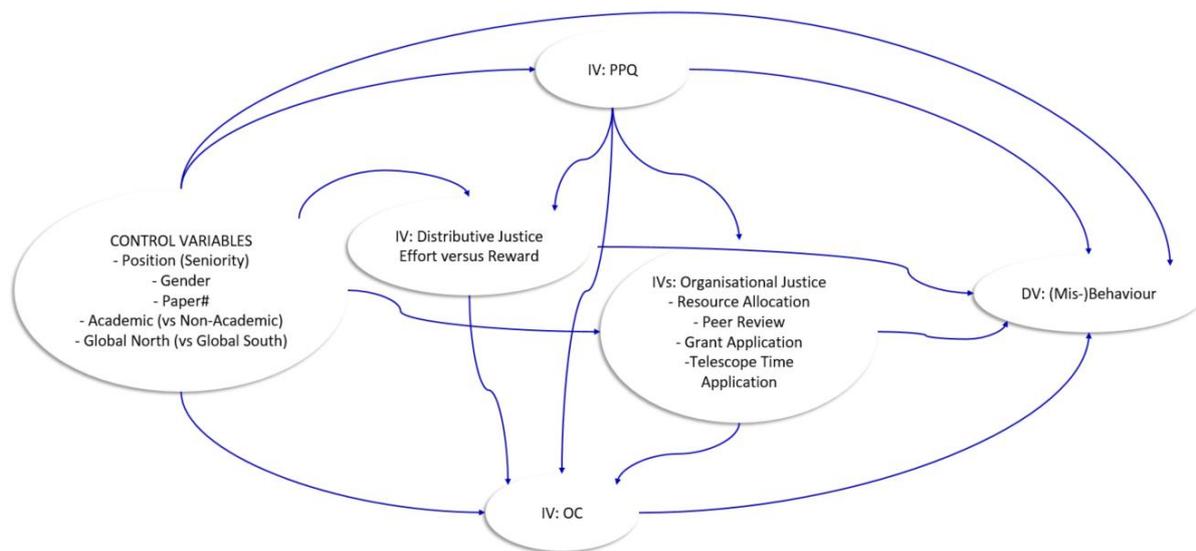


Figure 1. Visualisation of our SEM on the level of latent constructs of our control, independent (IV) and dependent variables (DV). Arrows indicate that a construct (arrowhead) is being regressed onto another construct (arrow start).

3.4. Statistical Analyses

The analysis of this survey was performed in SPSS and R. Data preparation, including recoding and calculation of mean scores, was performed in SPSS. We decided to exclude the 23 bachelor's and master's students from our sample, since we received too few responses from this category to conduct a proper analysis. All instruments measuring the independent variables (see S1), are scored on a scale from 1 (strongly disagree) to 5 (strongly agree) and are treated as continuous variables in our structural equation model (SEM). The steps to arrive at our final model started with testing the independent variable constructs PPQ and ERI (including overcommitment) by performing an exploratory factor analysis (EFA) for ordinal data (CATPCA in SPSS), and we derived Cronbach Alphas as scale reliabilities. For the EFA we used Promax Kaiser normalisation for rotating the factors. Next, all independent variable constructs were tested by means of confirmatory factor analysis (CFA) using Lavaan version 0.5-23 [53] in R version 3.3.1. For each construct, we used the residual correlation matrices to determine significant correlations of the indicators and included them into the respective models. After checking for construct validity, we further used Lavaan to perform structural equation modelling, which is the purpose it was designed for. Lavaan uses maximum likelihood estimation for regression analysis and listwise exclusion for missing data. The Results Section presents the results of our EFAs, CFAs and the complete SEM.

4. Results

4.1. Descriptive Statistics

In Table 1, we first present the descriptive statistics of the control variables. Females make up around 26% of the sample (N = 1827). For further analysis we combined the female and non-binary categories. Out of 2188 astronomers who shared their academic position in the survey, there are about 15% PhD candidates, 23% postdocs, 8.5% assistant professors, 26% full professors and 12% unranked astronomers. Out of the 2478 astronomers who declared whether their primary employment is at an academic university setting, 84% are employed in such setting and 16% are employed at other institutions that do research, such as national research institutes, observatories/telescopes or space agencies. Out of 1624 astronomers who answered in which country their primary employment is located, 85% astronomers work for an institution in the Global North and 15% in the Global South. In total, 2610 astronomers answered how many papers they published as first or co-authors in the last 5 years. The 8% who have not published any yet were excluded from our regression analysis, since they did not receive the item battery regarding organisational justice in terms of peer review. The largest publication category is 1–5 papers published in the last 5 years with 32% of respondents, followed by the 11–20 publications category with 16% of respondents. Only, 3.53% have been first or co-author for more than 100 papers in that time frame. For the subsequent analysis, we divided the number of published papers into three categories, in such a way that each group contains similar frequencies: 1–5 (828 cases), 6–20 (812 cases) and >20 papers (766 cases). Out of 2647 astronomers who answered the question of whether they applied for telescope time in the past 5 years, 59% replied yes and the rest replied no. The same amount of people filled in the question about whether they applied for a grant application in the past 5 years, and here 62% answered yes. Those who answered no for any of the two questions did not receive the item batteries regarding procedural justice with respect to telescope time or grant application processes, respectively.

Table 1. Descriptive statistics of the control variables of N = 3509 survey respondents. “Percent” is based on the total N (3509), while “Valid Percent” is based on “n valid”, which excludes missing data.

Baseline Characteristic	n Valid	Frequency	Percent	Valid Percent
Gender	1827			
Male		1333	38	73.0
Female		482	13.7	26.4
Non-Binary		12	0.3	0.7
Academic Position	2188			
PhD candidate		332	9.5	15.2
Postdoc/Research associate		504	14.4	23.0
Assistant professor		186	5.3	8.5
Associate professor		330	9.4	15.1
Full professor		568	16.2	26.0
Other		268	7.6	12.2
Academic/Non-Academic	2478			
Academic		2086	59.4	84.2
Non-Academic		392	11.2	15.8
Location of Employment	1624			
Global North		1377	39.2	84.8
Global South		247	7	15.2
Numbers of published papers (as first or co-author during the past 5 years)	2610			
1st paper currently under review		54	1.5	2.1
0		150	4.3	5.7
1–5		828	23.6	31.7
6–10		389	11.1	14.9

Table 1. Cont.

Baseline Characteristic	n Valid	Frequency	Percent	Valid Percent
11–20		423	12.1	16.2
21–30		173	4.9	6.6
31–40		158	4.5	6.1
41–50		131	3.7	5.0
51–60		61	1.7	2.3
61–70		43	1.2	1.6
71–80		40	1.1	1.5
81–90		31	0.9	1.2
91–100		37	1.1	1.4
>100		92	2.6	3.5

For each independent and dependent variable construct, we calculated the mean scores, which are presented in Table 2. The mean of the perceived publication pressure lies slightly above the mean of the scale. The effort versus reward ratio is 1.15, which means that the perceived effort put into work is higher than the perceived reward received for work. Astronomers also feel a slight overcommitment to work ($M = 3.39$). The four forms of organisational justice are generally above the mean of the scale, which indicates that astronomers tend to feel more justice than injustice when it comes to resource allocation, peer review, grant application and telescope time application. The mean of the perceived frequency of scientific misconduct lies just below the mean of the scale ($M = 2.99$). The mean impact of the 18 different misbehaviours (listed in Table 2) on the validity of the findings at hand (Quality Criterion 1) and on the resulting paper's ability to convey the research appropriately (Quality Criterion 2) are around 3.3. The mean value for the impact of misbehaviour on research diversity is higher ($M = 3.74$). However, one needs to consider that this question was only asked for the two types of misbehaviour (Item 8 and Item 9) for which it was expected that they have an impact on Quality Criterion 3. That choice was made in order not to burden the participants with a question that did not fit with the rest of the misbehaviour items.

Table 2. Means and standard deviations of the independent and dependent variables.

	n Valid	Mean	SD
Independent Variables			
Publication Pressure (PPQ)	1949	3.16	0.77
Distributive Justice: Effort	1757	3.68	0.86
Distributive Justice: Reward	1752	3.21	0.84
Organisational Justice: Resource Allocation	1852	3.02	0.8
Organisational Justice: Peer Review	1967	3.69	0.74
Organisational Justice: Grant Application	1267	3.14	0.79
Organisational Justice: Telescope Time Application	1185	3.36	0.72
Overcommitment	1755	3.39	0.84
Dependent Variables			
Occurrence of Misconduct	1869	2.99	0.66
Impact on Quality Criterion 1	1868	3.26	0.62
Impact on Quality Criterion 2	1868	3.29	0.64
Impact on Quality Criterion 3	1902	3.74	0.93

4.2. Exploratory Factor Analyses

In order to build our SEM, we first performed an EFA for the independent variables PPQ and ERI, the latter of which includes the overcommitment items. The Pearson correlation matrix for the PPQ items show considerably low correlations for 10 items ($<|0.3|$). When testing the Cronbach alphas (see Table S1a in S2) for the whole construct we found that removing those items would increase the reliability of the PPQ construct. We subsequently decided to remove those items, resulting in 12 remaining items. The Cronbach

alpha for the remaining PPQ construct is 0.871 (see Table S1b in S2), which is considered as good internal consistency. We henceforth used this cleaned PPQ for any further analysis. The CATPCA resulted in three factors, which we classified as (F1) extent and consequences on one's own conduct of research, (F2) impact on relationship with colleagues and (F3) suspected consequences on science (see Table S1c in S2).

The EFA for the ERI construct resulted in five factors (see Table S2a in S2): one factor representing the perceived effort put into the work (F3), one factor representing the perceived overcommitment to work (F5) and three factors representing the perception of being rewarded for one's work (F1: job situation, F2: salary, F4: receiving praise/respect). The Cronbach alphas are 0.691 for the effort construct, 0.780 for the overcommitment construct and 0.805 for the combined reward construct (see Table S2b–d in S2), which are considered acceptable values for the former two and good for the last one.

4.3. Confirmatory Factor Analyses

We subsequently ran CFAs for all independent variables, The results are presented in Table S3 in S2. This table includes the model fit indices CFI and TLI, where >0.9 indicates a good fit for both and RMSEA, where <0.05 denotes a good fit. In addition, the fourth column includes the Chi-square values for the difference between the models with and without accounting for significant covariation between indicators measuring the respective independent variable. All independent variable constructs show a good fit according to CFI and TLI. As for RMSEA, the fit is good for ERI and acceptable for the other constructs. For each independent variable, the model is unsurprisingly better when significant covariations between indicators are taken into account.

4.4. Structural Equation Model

This section presents the statistically significant main effects from the regression analysis of the whole structural equation model (Figure 1; $N = 520$ after listwise exclusion). The SEM fit is acceptable with a CFI of 0.802, a TLI of 0.793 and a RMSEA of 0.046, with a 90% CI (0.045, 0.047). For the sake of readability, we split the output by the independent and dependent variables, resulting in five different tables (see Tables 3–7; as a reminder, reference categories for the control variables are: gender: female/ non-binary; academic position: full professor; institute location: Global South; number of papers published: 1–5).

Table 3. Extracted output for the independent variable “Publication Pressure” from the regression analysis of our SEM model. * indicates statistical significance ($p < 0.05$).

IV: Publication Pressure	Estimate	Std.Err	z-Value	$p (> z)$	Std.lv	Std.all
Intercept:	3.159					
Gender: Male	−0.186	0.07	−2.64	0.008 *	−0.278	−0.124
Position: PhD	0.387	0.175	2.211	0.027 *	0.58	0.106
Position: Postdoc	0.506	0.091	5.537	<0.001 *	0.757	0.313
Position: Assistant Prof.	0.413	0.108	3.819	<0.001 *	0.618	0.194
Position: Associate Prof.	0.052	0.091	0.576	0.565	0.078	0.03
Position: Other	0.251	0.104	2.415	0.016 *	0.376	0.128
Primary Employer: Academic	0.052	0.121	0.426	0.67	0.077	0.02
Location: Global North	−0.357	0.094	−3.809	<0.001 *	−0.534	−0.181
Papers published: 6–20	−0.013	0.089	−0.147	0.883	−0.02	−0.009
Papers published: >20	−0.182	0.086	−2.109	0.035 *	−0.272	−0.136

Table 4. Extracted output for the two forms of the independent variable “Distributive Justice” from the regression analysis of our SEM model. * indicates statistical significance ($p < 0.05$).

IV: Reward (ERI)	Estimate	Std.Err	z-Value	$p (> z)$	Std.lv	Std.all
Intercept:	3.208					
PPQ	−0.385	0.072	−5.377	<0.001 *	−0.36	−0.36
Gender: Male	0.074	0.078	0.947	0.344	0.103	0.046
Position: PhD	−0.165	0.194	−0.851	0.395	−0.23	−0.042
Position: Postdoc	−0.32	0.102	−3.134	0.002 *	−0.447	−0.185
Position: Assistant Prof.	−0.176	0.119	−1.473	0.141	−0.245	−0.077
Position: Associate Prof.	−0.323	0.102	−3.169	0.002 *	−0.451	−0.171
Position: Other	−0.123	0.115	−1.071	0.284	−0.172	−0.058
Primary Employer: Academic	−0.015	0.134	−0.111	0.912	−0.021	−0.005
Location: Global North	−0.116	0.103	−1.126	0.26	−0.163	−0.055
Papers published: 6–20	0.225	0.1	2.261	0.024 *	0.314	0.144
Papers published: >20	0.248	0.096	2.57	0.01 *	0.346	0.173
IV: Effort (ERI)						
Intercept:	3.556					
PPQ	0.476	0.079	6.027	<0.001 *	0.414	0.414
Gender: Male	−0.226	0.088	−2.575	0.01 *	−0.294	−0.131
Position: PhD	−0.217	0.218	−0.999	0.318	−0.283	−0.052
Position: Postdoc	−0.277	0.113	−2.449	0.014 *	−0.361	−0.149
Position: Assistant Prof.	−0.121	0.134	−0.901	0.368	−0.157	−0.049
Position: Associate Prof.	−0.011	0.113	−0.095	0.924	−0.014	−0.005
Position: Other	−0.292	0.129	−2.253	0.024 *	−0.38	−0.129
Primary Employer: Academic	−0.049	0.151	−0.324	0.746	−0.064	−0.017
Location: Global North	0.174	0.116	1.501	0.133	0.227	0.077
Papers published: 6–20	−0.092	0.111	−0.83	0.406	−0.12	−0.055
Papers published: >20	0.069	0.107	0.644	0.52	0.09	0.045

Table 5. Extracted output for the four forms of the independent variable “Organisational Justice” from the regression analysis of our SEM model. * indicates statistical significance ($p < 0.05$).

IV: Organisational Justice (Resource Allocation)	Estimate	Std.Err	z-Value	$p (> z)$	Std.lv	Std.all
Intercept:	3.02					
PPQ	−0.274	0.057	−4.797	<0.001 *	−0.298	−0.298
Gender: Male	0.118	0.062	1.883	0.06	0.191	0.085
Position: PhD	0.359	0.156	2.294	0.022 *	0.582	0.106
Position: Postdoc	0.198	0.081	2.439	0.015 *	0.322	0.133
Position: Assistant Prof.	0.003	0.095	0.036	0.972	0.005	0.002
Position: Associate Prof.	−0.111	0.08	−1.391	0.164	−0.181	−0.068
Position: Other	0.071	0.092	0.773	0.44	0.115	0.039
Primary Employer: Academic	0.01	0.107	0.09	0.929	0.016	0.004
Location: Global North	0.072	0.082	0.877	0.38	0.117	0.04
Papers published: 6–20	0.15	0.079	1.896	0.058	0.244	0.112
Papers published: >20	0.088	0.076	1.151	0.25	0.142	0.071
IV: Organisational Justice (Peer Review)						
Intercept:	3.692					
PPQ	−0.418	0.074	−5.659	<0.001 *	−0.338	−0.338
Gender: Male	0.034	0.082	0.414	0.679	0.041	0.018
Position: PhD	0.526	0.206	2.557	0.011 *	0.638	0.117
Position: Postdoc	0.23	0.107	2.159	0.031 *	0.279	0.115
Position: Assistant Prof.	0.214	0.126	1.696	0.09	0.26	0.082
Position: Associate Prof.	0.066	0.106	0.62	0.535	0.08	0.03
Position: Other	0.152	0.122	1.245	0.213	0.184	0.062
Primary Employer: Academic	0.088	0.142	0.618	0.536	0.106	0.028
Location: Global North	−0.447	0.111	−4.041	<0.001 *	−0.542	−0.184
Papers published: 6–20	0.324	0.105	3.083	0.002 *	0.393	0.18
Papers published: >20	0.241	0.101	2.376	0.017 *	0.292	0.146

Table 5. Cont.

IV: Organisational Justice (Resource Allocation)	Estimate	Std.Err	z-Value	p (> z)	Std.lv	Std.all
IV: Organisational Justice (Grant Application)	Estimate	Std.Err	z-Value	p (> z)	Std.lv	Std.all
Intercept:	3.143					
PPQ	−0.473	0.081	−5.818	<0.001 *	−0.354	−0.354
Gender: Male	0.08	0.09	0.891	0.373	0.089	0.04
Position: PhD	0.648	0.225	2.886	0.004 *	0.725	0.132
Position: Postdoc	0.175	0.116	1.506	0.132	0.195	0.081
Position: Assistant Prof.	0.175	0.138	1.272	0.203	0.196	0.062
Position: Associate Prof.	−0.183	0.116	−1.579	0.114	−0.204	−0.077
Position: Other	0.128	0.133	0.966	0.334	0.143	0.049
Primary Employer: Academic	0.097	0.155	0.63	0.529	0.109	0.029
Location: Global North	−0.262	0.12	−2.194	0.028 *	−0.293	−0.099
Papers published: 6–20	0.202	0.114	1.769	0.077	0.226	0.103
Papers published: >20	0.193	0.11	1.749	0.08	0.215	0.108
IV: Organisational Justice (Telescope Time Application)						
Intercept:	3.356					
PPQ	−0.298	0.069	−4.303	<0.001 *	−0.255	−0.255
Gender: Male	0.121	0.081	1.501	0.133	0.155	0.069
Position: PhD	0.078	0.2	0.391	0.696	0.1	0.018
Position: Postdoc	0.085	0.104	0.822	0.411	0.109	0.045
Position: Assistant Prof.	0.002	0.123	0.016	0.987	0.003	0.001
Position: Associate Prof.	0.003	0.104	0.031	0.975	0.004	0.002
Position: Other	0.081	0.119	0.678	0.498	0.103	0.035
Primary Employer: Academic	−0.072	0.139	−0.515	0.606	−0.092	−0.024
Location: Global North	−0.215	0.107	−1.999	0.046 *	−0.274	−0.093
Papers published: 6–20	0.146	0.102	1.429	0.153	0.187	0.086
Papers published: >20	0.106	0.099	1.076	0.282	0.136	0.068

Table 6. Extracted output for the independent variable “Overcommitment” from the regression analysis of our SEM model.

* indicates statistical significance ($p < 0.05$).

IV: Overcommitment (ERI)	Estimate	Std.Err	z-Value	p (> z)	Std.lv	Std.all
Intercept:	3.433					
Publication Pressure	0.117	0.059	1.978	0.048 *	0.134	0.134
Reward (ERI)	−0.071	0.074	−0.964	0.335	−0.087	−0.087
Effort (RI)	0.459	0.059	7.714	<0.001 *	0.602	0.602
Organisational Justice (RA)	−0.039	0.072	−0.55	0.582	−0.041	−0.041
Organisational Justice (PR)	0.062	0.038	1.654	0.098	0.088	0.088
Organisational Justice (GA)	−0.001	0.037	−0.026	0.98	−0.001	−0.001
Organisational Justice (TA)	−0.068	0.041	−1.657	0.097	−0.091	−0.091
Gender: Male	0.006	0.056	0.098	0.922	0.009	0.004
Position: PhD	0.153	0.145	1.055	0.292	0.262	0.048
Position: Postdoc	0.042	0.081	0.516	0.606	0.072	0.03
Position: Assistant Prof.	−0.076	0.087	−0.878	0.38	−0.13	−0.041
Position: Associate Prof.	−0.164	0.074	−2.203	0.028 *	−0.28	−0.106
Position: Other	−0.01	0.084	−0.121	0.903	−0.017	−0.006
Primary Employer: Academic	−0.036	0.096	−0.379	0.705	−0.062	−0.016
Location: Global North	0.043	0.077	0.555	0.579	0.074	0.025
Papers published: 6–20	0.006	0.072	0.087	0.931	0.011	0.005
Papers published: >20	0.003	0.069	0.048	0.962	0.006	0.003

Table 7. Extracted output for the dependent variable “Frequency of Misbehaviour Occurrence” from the regression analysis of our SEM model. * indicates statistical significance ($p < 0.05$).

DV: Frequency of Misbehaviour Occurrence	Estimate	Std.Err	z-Value	$p (> z)$	Std.lv	Std.all
Intercept:	2.992					
Publication Pressure	0.373	0.066	5.679	<0.001 *	0.433	0.433
Reward (ERI)	−0.138	0.072	−1.932	0.053	−0.172	−0.172
Effort (RI)	−0.021	0.057	−0.366	0.714	−0.028	−0.028
Organisational Justice (RA)	0.09	0.068	1.321	0.187	0.096	0.096
Organisational Justice (PR)	−0.062	0.036	−1.733	0.083	−0.089	−0.089
Organisational Justice (GA)	−0.004	0.035	−0.112	0.911	−0.006	−0.006
Organisational Justice (TA)	−0.106	0.04	−2.673	0.008 *	−0.144	−0.144
Overcommitment (ERI)	0.073	0.072	1.022	0.307	0.074	0.074
Gender: Male	−0.06	0.053	−1.13	0.259	−0.104	−0.046
Position: PhD	−0.062	0.137	−0.455	0.649	−0.108	−0.02
Position: Postdoc	0.014	0.077	0.186	0.853	0.025	0.01
Position: Assistant Prof.	0.052	0.082	0.641	0.522	0.091	0.029
Position: Associate Prof.	−0.023	0.07	−0.325	0.745	−0.04	−0.015
Position: Other	−0.054	0.079	−0.677	0.498	−0.093	−0.032
Primary Employer: Academic	−0.158	0.09	−1.752	0.08	−0.275	−0.072
Location: Global North	0.205	0.074	2.762	0.006 *	0.355	0.121
Papers published: 6–20	0.005	0.067	0.071	0.943	0.008	0.004
Papers published: >20	0.074	0.065	1.137	0.255	0.129	0.065

Table 3 presents the main effects of the control variables regressed onto perceived publication pressure. Being male as opposed to female/non-binary decreases the chance to perceive publication pressure by 0.186 points. Astronomers occupying a position other than associate professor tend to feel more publication pressure than a full professor. Whether one works at an academic institution or not does not have a significant effect on publication pressure. However, those working at an institution located in the Global North are less likely to feel publication pressure by 0.357 points. Having published more than 20 papers in the last 5 years decreases the likelihood of perceiving publication pressure as compared to having published between 1 and 5 papers.

As for the distributive justice factors reward and effort (Table 4), astronomers who perceive publication pressure feel less rewarded (by 0.385 points) for the work they do, while at the same time they feel that they put more effort (by 0.476 points) into their work than astronomers who feel less publication pressure. Males tend to feel like they need to put less effort into their work than females or non-binaries (by 0.226 points). Postdocs tend to feel less rewarded for their work by 0.32 points, while at the same time also putting less effort than full professors by 0.277 points. Associate professors also feel that they are less rewarded as compared to full professors by 0.323 points, but they show no significant effect on the effort factor. Neither being an academic astronomer, nor being employed in the Global North makes a difference in the reward and effort factors as compared to the opposite. Astronomers who have published more than five papers feel more rewarded for their job than those who have published less.

The perception of all four kinds of organisational justice measured depends on the perceived publication pressure (Table 5). The feeling of being treated fairly decreases with increasing publication pressure for all four latent variables (with parameter estimates between 0.274 and 0.473). PhDs and postdocs feel treated more fairly in terms of resource allocation and peer review than full professors. PhDs also feel more justice when it comes to grant application processes. Being academic or not does not have a significant effect on any of the four organisational justice perceptions. Being employed in the Global North decreases the likelihood of perceiving fairness in terms of peer review and grant and telescope applications as compared to those employed in the Global South. Those who have published more than five papers feel more organisational justice in terms of peer review than the reference category (1–5 papers).

The feeling of being overcommitted to work unsurprisingly significantly increases with increasing perceived publication pressure (0.117 points) and with perceived effort put into work (0.459 points; Table 6). Associate professors tend to feel less overcommitted to their work compared to full professors by 0.164 points.

Finally, we turn to the results of our dependent variable: perception of how often misconduct occurs in astronomy (Table 7). The parameter estimate of the main effect of publication pressure on perception of misconduct frequency is 0.373, which means that increasing perceived publication pressure leads to increased perception of misconduct. Perceived fairness of telescope time application processes also has a significant effect on the perception of misconduct; a decreasing feeling of fairness increases the perception of misconduct by 0.106 points. Being employed in the Global North also increases the perception of misconduct by 0.205 points. In addition to the main effects on perceived frequency of misconduct, our model calculates the mediated effects. Let us first attend the results of the effects of the control variables as mediated by publication pressure. Being male as compared to female/non-binary reduces the perception of misconduct by 0.069 points. Any other position than associate professor increases the chance to perceive misconduct as compared to full professor. Employment at an institution in the Global North decreases the perception of misconduct by 0.133 points, mediated through the perception publication pressure. Having published more than 20 papers also decreases the likelihood of perceiving misconduct. Publication pressure, as mediated through organisational justice in terms of telescope time application increases the likelihood of perceiving misconduct by 0.032 points.

4.5. Perceived Impact on Research Quality

Lastly, we analysed the perceived impact of the 18 types of misbehaviour on the three aspects of research quality: the validity of the findings at hand (Quality Criterion 1; QC1), the resulting paper's ability to convey the research appropriately (Quality Criterion 2; QC2) and the impact on research diversity (Quality Criterion 3; QC3). Following Bouter et al. (2016) and Haven et al. (under review), we calculated the perceived impact as the product score of the means of perceived frequency and impact on the respective quality criterion (means are listed in Table S1 in S3). The higher the resulting number, the higher the harm on research. Table S2 in S3 lists the results of perceived impact scores in descending order. In addition, Figure 2 visualises the perceived impact on research quality in four quadrants, indicating high frequency and high impact (Q1), low frequency and low impact (Q3), low frequency and high impact (Q2) and high frequency and low impact (Q4). Item 8 ("Propose study questions solely because they are considered a 'hot' topic") and Item 9 ("Not considering a study question because it is not considered a 'hot' topic, even though it could be important for astronomy") show the highest perceived impact on research quality (14.33 and 12.89, respectively). This impact is related to QC3 ("impact on research diversity") and can be found in Q1. While the high impact of these types of misbehaviour on QC3 ($M = 3.7$ and $M = 3.79$ for Item 8 and 9, respectively) may well be expected, the comparatively high frequencies ($M = 3.88$ and $M = 3.4$ for Item 8 and 9, respectively) are an interesting result. Item 18 ("Biased interpretation of data that distorts results") follows, with perceived impact amounting to 12.47 for QC1 and 12.17 for QC2. As we can see from Figure 2, these high values for perceived impact stem mostly from the comparatively high impacts on the two quality criteria ($M = 4.04$ and $M = 3.95$), rather than a high frequency ($M = 3.08$), which is just above the mean (2.99). The occurrence of Item 13 ("Data fabrication and/or falsification") also has a high impact on both quality criteria ($M = 4.17$ for QC1, $M = 4.01$ for QC2). Due to the comparative low frequency of Item 13 ($M = 1.95$), the perceived impact of this type of misbehaviour corresponds to rank 30 (QC1) and 31 (QC2) out of 38. Item 13 can be found in Q2. Item 10 ("Giving authorship credit to someone who has not contributed substantively to a manuscript"), which is ranked lower than Item 13 (32 for QC2 and 35 for QC1), can be interpreted as the opposite of Item 13, in that Item 10 has a high occurrence ($M = 3.73$) but a low impact on quality criteria

($M = 1.97$ for QC1 and $M = 2.06$ for QC2). Item 10 can therefore be found in Q4. By contrast, “Denying authorship credit to someone who has contributed substantively to a manuscript” (Item 11) is located in Q3, since it does not occur very often ($M = 2.05$) and when it occurs it also has a comparatively low impact on QC1 ($M = 2.53$) and QC2 ($M = 2.67$). The perceived impact of this type of misbehaviour is also the lowest ranked. Plagiarism, a severe form of scientific misconduct, is represented by Items 15 and 16 (“Using published ideas or phrases of others without referencing (Plagiarism)” and “Using unpublished ideas or phrases of others without their permission”, respectively). Both are also located in Q3, due to their rather low occurrence ($M = 2.61$ and $M = 2.41$, respectively) and only low-to-moderate impact on QC1 ($M = 2.82$ and $M = 2.80$, respectively) and QC2 ($M = 3.18$ and $M = 3.07$, respectively). This puts Item 15 on rank 29 (QC2) and rank 34 (QC1) and Item 16 on rank 33 (QC2) and rank 36 (QC1).

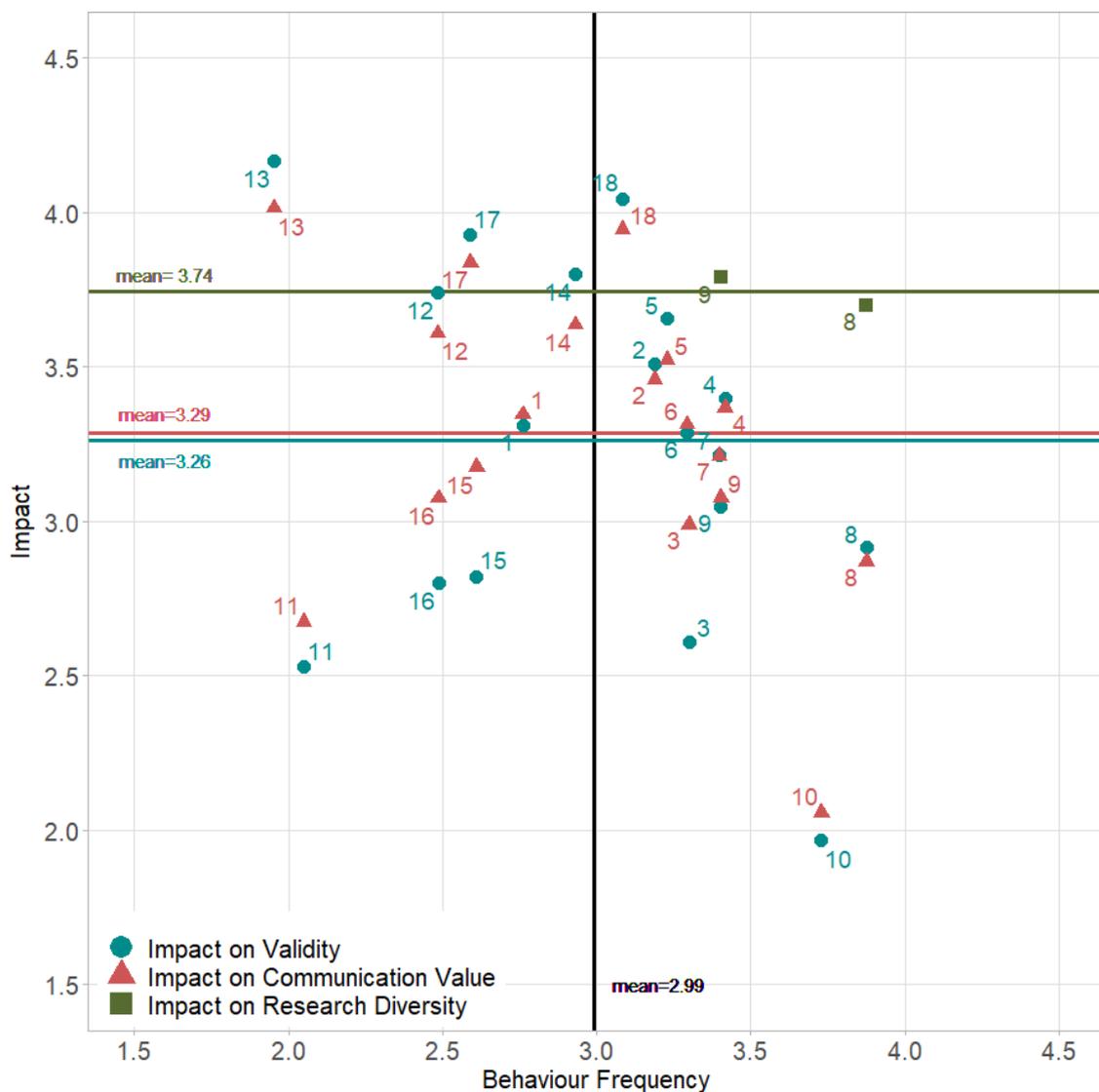


Figure 2. Impact on the three aspects of research quality for each type of misbehaviour versus the frequency of occurrence of each type of misbehaviour. Circles denote QC1, triangles refer to QC2 and the square to QC3. The four quadrants indicate high frequency and high impact (Q1), low frequency and low impact (Q3), low frequency and high impact (Q2) and high frequency and low impact (Q4).

5. Discussion

Building on previous quantitative research on scientific misconduct [17,19,24–26,51,54] and qualitative research on deviant behaviour in astronomy [10,20], we built a structural equation model relating role-associated factors, such as academic position and location of employment with environmental factors, such as perceived publication pressure and distributive and organisational justice, as well as our dependent variables of scientific misconduct and research quality. We found that the location of the institution where an astronomer is employed in terms of the Global North versus Global South makes up about 5% of the variance of observed misconduct. Perceived organisational justice in terms of telescope time application processes explains 3% and perceived publication pressure explains nearly 10% of the variance of observed misconduct.

In addition of publication pressure having a direct effect on scientific misconduct (cf. [36]), we worked from the assumption that perceived publication pressure influences the perception of distributive and organisational justice, which our results confirm. An astronomer who perceives publication pressure is more likely to perceive less reward from their work, less organisational justice (in terms of resource allocation, peer review, grant application and telescope time application), the need to put more effort into work and more overcommitment to work. Hence, publication pressure is indeed a key factor to determine how research culture and integrity in astronomy is perceived.

Publication pressure, in turn, is more likely to be perceived by astronomers with academic ranks below a full professor (cf. [31,34]). Interestingly, astronomers employed at institutions in the Global North feel less publication pressure, despite observing more misconduct. Hence, there is some difference between institutions in the Global North as compared to the Global South, which makes astronomers perceive more publication pressure (which in turn increases the likelihood of observing misconduct), yet at the same time suppresses the perception of scientific misconduct. Unsurprisingly, there is a tendency to perceive less publication pressure when one has published more than 20 first or co-authored papers in the last 5 years. Males do not only perceive less publication pressure than females/non-binaries, but also feel like they need to put less effort into their work. This perception is consistent with research showing that working conditions are harder for females than males (For a good overview of such studies refer to: <https://theconversation.com/amp/looking-at-the-stars-or-falling-by-the-wayside-how-astronomy-is-failing-female-scientists-159139>; accessed on 18 June 2021).

Early-career researchers perceive more organisational justice in terms of resource allocation, peer review and grant allocation than full professors. This may be because early-career researchers may still have a positive opinion of organisational processes, whereas more experience may lead to more occasions of unfairness (cf. [10]).

The frequency of observed misbehaviour shows that the severe types of misbehaviour (FFP type) such as data fabrication and falsification (Item 13), concealing results (Item 17) and forms of plagiarism (Items 15 and 16) occur less often than the QRPs, as expected [17,54]. Among the most frequently occurring QRPs are making the topic selection too dependent on whether a topic is or is not “hot” (Items 8 and 9), questionable authorship practices (Item 10) and insufficient supervision (Item 4). “Biased interpretation of data that distorts results” (Item 18) occurs a less often than Item 10; however, it is ranked among the most severe QRPs in terms of impact on research quality, as opposed to Item 10. Our findings agree with Bouter et al. [51], who found that data fabrication and falsification (Item 13) is believed to be the biggest threat to the validity of the findings (QC1) at hand and the communication value of the resulting paper (QC2). In comparison, plagiarism (an FFP-type misbehaviour) is ranked very low on the impact on QC1 and QC2 scores. This makes sense, since as Haven et al. [36] point out, “plagiarism fails to connect the knowledge to its proper origin, but it need not distort scientific knowledge per se”, whereas falsification and fabrication do. Our perceived impact ranking (the product of the mean scores of perceived frequency and impact on research quality) also agrees with the findings of Bouter et al. [51]. It suggests that is not outright fraud that dominates the negative impact on knowledge

creation, but rather behaviour that cuts corners to publishable output (Item 8, 9 and 5). We conclude that many a little makes a mickle; the epistemic harm in research in astronomy done by QRPs seems to be greater than that done by FFPs, which would agree with finding by Haven [54] and Bouter et al. [51].

6. Strengths/Limitations

Surveys and data analysis come with strengths and limitations. Let us first turn to the strengths of this study. First, we sampled astronomers from all over the world. This paper gives a snapshot of the international cultural climate in astronomy and its impact on scientific misconduct and quality, while previous studies on this topic mainly focused on universities in the US or the Netherlands (e.g., [23,27]). The second strong point of this survey is that the types of misconduct we chose are not only based on previous literature on this topic (e.g., [17,19,24–26,54]), but also on qualitative research conducted on deviant behaviour in the field of astronomy [10]. As Hesselmann [1] (p. 61f.) points out, “the meaning of misbehaviour is permanently shifting” and therefore, “measuring scientific misconduct quantitatively should not be first on our research agenda.” We believe that the qualitative study by Heuritsch [10] gave us solid ground to tailor the quantitative studies performed on misbehaviour to the field of astronomy. Third, our analysis is the first among the literature relating publication pressure and distributive and organisational justice with scientific misconduct, using structural equation modelling, allowing for an estimation of the model in its whole complexity. Fourth, it is also the first study in this set of literature that operationalises research quality. We therefore measure the impact of scientific misconduct on scientific quality instead of implying that relationship through the concept of research integrity.

Our study also comes with several limitations. First, while our response rate was acceptable (25%), our completion rate lies at around 14%, which may be considered as relatively low, but can still be compared to that of similar web-based surveys (e.g., [27]). The reason for this drop may be that the survey was considered long, as evident from some feedback from the respondents; it took 30–60 min to complete. Second, due to the length of the survey, we may need to consider response bias towards those who feel publication pressure and feel unfairly treated and, hence, may be more enthusiastic in voicing their opinion about this topic. On the one hand, this may overestimate the effects of publication pressure and organisational injustice on misconduct. On the other hand, those who left the field of astronomy as a result of publication pressure and injustice are of course not sampled; hence, publication pressure and organisational injustice may be underestimated through survivor bias [55]. Third, respondents criticised that there was no “NA” option for the experience of misbehaviour items. While respondents did not have to choose an answer to move forward in the survey, this may have resulted in an underestimation of occurrence of misconduct, as astronomers who may not have much experience in the field, clicked the lowest or the middle answer category, while having preferred an NA option. Because self-reports may result in underreporting misbehaviour [1], we chose to ask about the general experience with the types of misbehaviour. Therefore, we expect that the underreporting of one’s own misconduct would mitigate the overreporting of misconduct by others [36]. Fourth, many filter questions, such as if one has already applied for grants or telescope time, resulted in a comparatively small sample (N = 520) for the SEM analysis, because of list-wise exclusion. Fifth, at the time we designed the survey we had no knowledge about the revised PPQ [37], which we would have used instead of the PPQ and may have resulted into better construct validity, without having to adapt the construct for our own further analysis. Sixth, while it was our theoretical aim to conduct a census of all worldwide astronomers, this was practically impossible due to time, budget and resource constraints. Our three-stage sampling design aimed at completeness in getting a hold on astronomical institutions worldwide and reaching as many astronomers as possible. However, we cannot expect that our list is indeed complete, nor that our contacts reached all astronomers from the respective institutes, nor that our sample is random. Therefore,

representativeness may be limited. To improve this, further testing for item batteries may also inform weighting to adjust the sample proportions to the population proportion.

7. Conclusions, Implications and Outlook for Further Research

The aim of this research was to study the impact of perceptions of publication pressure and distributive and organisational justice on the observation of occurrence of scientific misconduct and the impact that certain types of misconduct have on research quality in astronomy. While we did not find statistically significant effects of perceived distributive and organisational justice of the four processes—resource allocation, peer review, grant application and telescope time application—on research misbehaviour, we strongly emphasise that publication pressure is part of research culture. As outlined by Heuritsch [10], institutional norms define what is seen as a good researcher, and publication rate is one of the key indicators to measure the performance of an astronomer. Arguably, playing the indicator game is an innovative path to success, so we worked from the assumption that research (mis-)behaviour is reflexive, insofar that it depends on how one's performance is evaluated. We found that publication pressure explains 10% of the variance of occurrence of misconduct and between 7% and 13% of the variance of the perception of distributive and organisational justice as well as overcommitment to work. We subsequently analysed the impact of the individual types of misbehaviour on three aspects of research quality. We agree with findings of previous studies (e.g., [51,54]) that QRPs should not be underestimated for their epistemic harm.

We conclude that there is a need for a policy change. In the distribution of institutional rewards, grants and telescope time, less attention to metrics (such as publication rate) would foster better scientific conduct and, hence, research quality. Publication pressure could also be reduced by reconsidering what is considered publishable. Since, for example, negative results cannot easily be published [10], a lot of scientific work may not be recognised as valuable research. Future studies could draw on astronomers' own visions of what is valuable about their research work and could work out more diverse ways of evaluation on that basis. This could include potentially exploring more innovative ways of creating quality output that may count towards one's performance. This requires reflecting and working on the structural conditions that comprise the norms and culture of research in astronomy. After all, they comprise the external constituents of the *situation* of an actor and are therefore of high relevance in the individual's actions.

Rational choice theory provided us with the appropriate framework, encompassing many partial theories regarding scientific misconduct used in earlier studies. Drawing on organisational culture theory and the NPM theory of explaining misconduct, we used constructs, such as perceived organisational justice and publication pressure, as independent variables predicting scientific misconduct and its impact on research quality. Future research could further build on these theories, for example, by using the SORC [22] to evaluate research climate or by studying effects of NPM practices other than measuring performance through the publication rate, using, for example, audits. While the bad apple theory may be too simplistic to stand alone, individual dispositions may regulate response to strain, making researchers more or less prone to (resist) engaging in misbehaviour. Future quantitative studies may complete the rational choice picture, by paying tribute to this internal component of the astronomers' research situation. Moreover, one could study their motivation to do research and to publish and could relate this to the importance of publications in the field.

Supplementary Materials: The followings are available online at <https://www.mdpi.com/article/10.3390/publications9040052/s1>. S1: Survey Questions, S2: EFAs and CFAs, S3: Perceived impact of scientific misconduct on research quality.

Funding: This study was performed in the framework of the junior research group "Reflexive Metrics", which is funded by the BMBF (German Bundesministerium für Bildung und Forschung; project number: 01PQ17002).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: First: I would like to extend my gratitude to the 3509 astronomers who dedicated upwards of half an hour—despite the publish-or-perish imperative—to participate in this survey. Second, Thea Gronemeier and Florian Beng assisted the survey design and were a big support in data processing. Third, I would like to thank my supervisor, Stephan Gauch, for facilitating this project. Fourth, thank you to all the pre-testers: Niels Taubert, Jens Ambrasat, Andrej Dvornik, Iva Laginja, Levente Borvák, Nathalie Schwichtenberg, Theresa Velden, Richard Heidler, Rudolf Albrecht, Andreas Herdin, Alex Fenton and Philipp Löschl. Fifth, I would like to thank Enrique Garcia Bourne for the mental support throughout, the proofreading and for assisting me with the script that helped me reach out to thousands of IAU members. Finally, I am grateful to GESIS (Lebniz Institut für Sozialwissenschaften) for the scholarship that enabled me to participate in their 2019 survey methodology course, where I met Sonila Dardha—an extraordinarily competent and kind survey methodologist, whose feedback was vital for this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Hesselmann, F.; Wienefoet, V.; Reinhart, M. Measuring Scientific Misconduct—Lessons from Criminology. *Publications* **2014**, *2*, 61–70. [[CrossRef](#)]
- Stephan, P. *How Economics Shapes Science*; Harvard University Press: Cambridge, MA, USA, 2012.
- Laudel, G.; Gläser, J. Beyond breakthrough research: Epistemic properties of research and their consequences for research funding. *Res. Policy* **2014**, *43*, 1204–1216. [[CrossRef](#)]
- Fochler, M.; De Rijcke, S. Implicated in the Indicator Game? An Experimental Debate. *Engag. Sci. Technol. Soc.* **2017**, *3*, 21–40. [[CrossRef](#)]
- Desrosières, A. *The Politics of Large Numbers—A History of Statistical Reasoning*; Harvard University Press: Cambridge, MA, USA, 1998; ISBN 9780674009691.
- Porter, T. *Trust in Numbers*; Princeton University Press: Princeton, NJ, USA, 1995.
- Dahler-Larsen, P. Constitutive Effects of Performance Indicators: Getting beyond unintended consequences. *Public Manag. Rev.* **2014**, *16*, 969–986. [[CrossRef](#)]
- Dahler-Larsen, P. *Quality—From Plato to Performance*; Palgrave Macmillan: Cham, Switzerland, 2019. [[CrossRef](#)]
- Wouters, P. Bridging the Evaluation Gap. *Engag. Sci. Technol. Soc.* **2017**, *3*, 108–118. [[CrossRef](#)]
- Heuritsch, J. The Evaluation Gap in Astronomy—Explained through a Rational Choice Framework. *arXiv* **2021**, arXiv:2101.03068.
- Lorenz, C. If You’re So Smart, Why Are You under Surveillance? Universities, Neoliberalism, and New Public Management. *Crit. Inq.* **2012**, *38*, 599–629. [[CrossRef](#)]
- Anderson, M.S.; Ronning, E.A.; De Vries, R.; Martinson, B.C. The Perverse Effects of Competition on Scientists’ Work and Relationships. *Sci. Eng. Ethics* **2007**, *13*, 437–461. [[CrossRef](#)]
- Tijdink, J.K.; Smulders, Y.M.; Vergouwen, A.C.M.; de Vet, H.C.W.; Knol, D.L. The assessment of publication pressure in medical science; validity and reliability of a Publication Pressure Questionnaire (PPQ). *Qual. Life Res.* **2014**, *23*, 2055–2062. [[CrossRef](#)]
- Moosa, I.A. *Publish or Perish—Perceived Benefits Versus Unintended Consequences*; Edward Elgar Publishing: Northampton, MA, USA, 2018. [[CrossRef](#)]
- Rushforth, A.D.; De Rijcke, S. Accounting for Impact? The Journal Impact Factor and the Making of Biomedical Research in the Netherlands. *Minerva* **2015**, *53*, 117–139. [[CrossRef](#)] [[PubMed](#)]
- U.S. Office of Science and Technology Policy, Executive Office of the President. OSTP, Federal Policy on Research Misconduct. 2000. Available online: http://www.Ostp.Gov/html/001207_3.Html (accessed on 10 September 2021).
- Martinson, B.C.; Anderson, M.S.; de Vries, R. Scientists behaving badly. *Nature* **2005**, *435*, 737–738. [[CrossRef](#)]
- Haven, T.L.; van Woudenberg, R. Explanations of Research Misconduct, and How They Hang Together. *J. Gen. Philos. Sci.* **2021**, *19*, 1–9.
- Martinson, B.C.; Anderson, M.S.; Crain, A.L.; De Vries, R. Scientists’ perceptions of organizational justice and self-reported misbehaviors. *J. Empir. Res. Hum. Res. Ethics* **2006**, *1*, 51–66. [[CrossRef](#)]
- Heuritsch, J. Effects of metrics in research evaluation on knowledge production in astronomy a case study on Evaluation Gap and Constitutive Effects. In Proceedings of the STS Conference Graz 2019, Graz, Austria, 6–7 May 2019. [[CrossRef](#)]
- Crain, L.A.; Martinson, B.C.; Thrush, C.R. Relationships between the Survey of Organizational Research Climate (SORC) and self-reported research practices. *Sci. Eng. Ethics* **2013**, *19*, 835–850. [[CrossRef](#)]
- Martinson, B.C.; Thrush, C.R.; Crain, A.L. Development and validation of the Survey of Organizational Research Climate (SORC). *Sci. Eng. Ethics* **2013**, *19*, 813–834. [[CrossRef](#)]
- Wells, J.A.; Thrush, C.R.; Martinson, B.C.; May, T.A.; Stickler, M.; Callahan, E.C. Survey of organizational research climates in three research intensive, doctoral granting universities. *J. Empir. Res. Hum. Res. Ethics* **2014**, *9*, 72–88. [[CrossRef](#)]

24. Martinson, B.C.; Nelson, D.; Hagel-Campbell, E.; Mohr, D.; Charns, M.P.; Bangerter, A. Initial results from the Survey of Organizational Research Climates (SOuRCe) in the U.S. department of veterans affairs healthcare system. *PLoS ONE* **2016**, *11*, e0151571. [[CrossRef](#)] [[PubMed](#)]
25. Martinson, B.C.; Crain, A.L.; Anderson, M.S.; De Vries, R. Institutions 'Expectations for Researchers' Self-Funding, Federal Grant Holding, and Private Industry Involvement: Manifold Drivers of Self-Interest and Researcher Behavior. *Acad. Med.* **2009**, *84*, 1491–1499. [[CrossRef](#)] [[PubMed](#)]
26. Martinson, B.C.; Crain, L.A.; De Vries, R.; Anderson, M.S. The importance of organizational justice in ensuring research integrity. *J. Empir. Res. Hum. Res. Ethics* **2010**, *5*, 67–83. [[CrossRef](#)] [[PubMed](#)]
27. Haven, T.L.; Bouter, L.M.; Smulders, Y.M.; Tijdink, J.K. Perceived publication pressure in Amsterdam—Survey of all disciplinary fields and academic ranks. *PLoS ONE* **2019**, *14*, e0217931. [[CrossRef](#)]
28. Zuiderwijk, A.; Spiers, H. Sharing and re-using open data: A case study of motivations in astrophysics. *Int. J. Inf. Manag.* **2019**, *49*, 228–241. [[CrossRef](#)]
29. Bedeian, A.; Taylor, S.; Miller, A. Management science on the credibility bubble: Cardinal sins and various misdemeanors. *Acad. Manag. Learn. Educ.* **2010**, *9*, 715–725.
30. Bouter, L.M. Commentary: Perverse incentives or rotten apples? *Acc. Res.* **2015**, *22*, 148–161. [[CrossRef](#)] [[PubMed](#)]
31. Tijdink, J.K.; Verbeke, R.; Smulders, Y.M. Publication pressure and scientific misconduct in medical scientists. *J. Empir. Res. Hum. Res. Ethics* **2014**, *9*, 64–71. [[CrossRef](#)] [[PubMed](#)]
32. Tijdink, J.K.; Vergouwen, A.C.M.; Smulders, Y.M. Publication pressure and burn out among Dutch medical professors: A nationwide survey. *PLoS ONE* **2013**, *3*, e73381. [[CrossRef](#)] [[PubMed](#)]
33. Tijdink, J.K.; Schipper, K.; Bouter, L.M.; Pont, P.M.; De Jonge, J.; Smulders, Y.M. How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers. *BMJ Open* **2016**, *6*, e008681. [[CrossRef](#)] [[PubMed](#)]
34. Miller, A.N.; Taylor, S.G.; Bedeian, A.G. Publish or perish: Academic life as management faculty live it. *Career Dev. Int.* **2011**, *16*, 422–445. [[CrossRef](#)]
35. Van Dalen, H.P.; Henkens, K. Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. *J. Am. Soc. Inf. Sci. Technol.* **2012**, *63*, 1282–1293. [[CrossRef](#)]
36. Haven, T.L.; Tijdink, J.K.; Martinson, B.C.; Bouter, L.; Oort, F. Explaining variance in perceived research misbehaviour. *Res. Integr. Peer Rev.* **2021**, *6*, 7. [[CrossRef](#)]
37. Haven, T.L.; Tijdink, J.K.; De Goede, M.E.E.; Oort, F. Personally perceived publication pressure—Revising the Publication Pressure Questionnaire (PPQ) by using work stress models. *Res. Integr. Peer Rev.* **2019**, *4*, 7. [[CrossRef](#)]
38. Sovacool, B.K. Exploring scientific misconduct: Isolated individuals, impure institutions, or an inevitable idiom of modern science? *J. Bioethical Inq.* **2008**, *5*, 271–282. [[CrossRef](#)]
39. Hackett, E.J. A Social Control Perspective on Scientific Misconduct. *J. High. Educ.* **1994**, *65*, 242–260. [[CrossRef](#)]
40. Agnew, R. Foundation for a general strain theory of crime and delinquency. *Criminology* **1992**, *30*, 47–87. [[CrossRef](#)]
41. Merton, R.K. Social structure and anomie. *Am. Sociol. Rev.* **1938**, *3*, 672–682. [[CrossRef](#)]
42. Espeland, W.N.; Vannebo, B. Accountability, Quantification, and Law. *Annu. Rev. Law Soc. Sci.* **2008**, *3*, 21–43. [[CrossRef](#)]
43. Halfman, W.; Radder, H. The Academic Manifesto: From an Occupied to a Public University. *Minerva* **2015**, *53*, 165–187. [[CrossRef](#)]
44. Overman, S.; Akkerman, A.; Torenvlied, R. Targets for honesty: How performance indicators shape integrity in Dutch higher education. *Public Adm.* **2016**, *94*, 1140–1154. [[CrossRef](#)]
45. Atkinson-Grosjean, J.; Fairley, C. Moral Economies in Science: From Ideal to Pragmatic. *Minerva* **2009**, *47*, 147–170. [[CrossRef](#)]
46. Esser, H. Soziologie. Spezielle Grundlagen. Band 1: Situationslogik und Handeln. *KZfSS Kölner Z. Soziologie Soz.* **1999**, *53*, 773. [[CrossRef](#)]
47. Coleman, J.S. *Foundations of Social Theory*; Belknap Press of Harvard University Press: Cambridge, MA, USA; London, UK, 1990.
48. Roy, J.R.; Mountain, M. The Evolving Sociology of Ground-Based Optical and Infrared astronomy at the Start of the 21st Century. In *Organizations and Strategies in Astronomy*; Heck, A., Ed.; Astrophysics and Space Science Library; Springer: Dordrecht, The Netherlands, 2006; Volume 6, pp. 11–37.
49. Chang, H.-W.; Huang, M.-H. The effects of research resources on international collaboration in the astronomy community. *J. Assoc. Inf. Sci. Technol.* **2015**, *67*, 2489–2510. [[CrossRef](#)]
50. Heidler, R. Cognitive and Social Structure of the Elite Collaboration Network of Astrophysics: A Case Study on Shifting Network Structures. *Minerva* **2011**, *49*, 461–488. [[CrossRef](#)]
51. Bouter, L.M.; Tijdink, J.; Axelsen, N.; Martinson, B.C.; ter Riet, G. Ranking major and minor research misbehaviors: Results from a survey among participants of four World Conferences on Research Integrity. *Res. Integr. Peer Rev.* **2016**, *1*, 17. [[CrossRef](#)] [[PubMed](#)]
52. Siegrist, J.; Li, J.; Montano, D. *Psychometric Properties of the Effort-Reward Imbalance Questionnaire*; Department of Medical Sociology, Faculty of Medicine, Duesseldorf University: Düsseldorf, Germany, 2014.
53. Rosseel, Y. lavaan: An R Package for Structural Equation Modeling. *J. Stat. Softw.* **2012**, *48*, 1–36. Available online: <https://www.jstatsoft.org/v48/i02/> (accessed on 10 September 2021). [[CrossRef](#)]

-
54. Haven, T.L. Towards a Responsible Research Climate: Findings from Academic Research in Amsterdam. 2021. Available online: <https://research.vu.nl/en/publications/towards-a-responsible-research-climate-findings-from-academic-res> (accessed on 10 September 2021).
 55. Kurtz, M.J.; Henneken, E.A. Measuring Metrics—A 40-Year Longitudinal Cross-Validation of Citations, Downloads, and Peer Review in Astrophysics. *J. Assoc. Inf. Sci. Technol.* **2017**, *68*, 695–708. [[CrossRef](#)]