DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Influencing factors to research integrity and research misconduct: case analysis and findings from Hungarian universities

by Anna Catharina Vieira Armond

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF HEALTH SCIENCES

DEBRECEN, 2022

DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Influencing factors to research integrity and research misconduct: case analysis and findings from Hungarian universities

by Anna Catharina Vieira Armond

Supervisor: Dr. Péter Kakuk

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF HEALTH SCIENCES

DEBRECEN, 2022

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION AND DISSERTATION BACKGROUND 5
Introduction 5
Literature review 5
Concepts and definitions 5
Consequences of research misbehavior 7
Causal factors of research misbehavior 8
Beyond research misconduct 9
Fostering research integrity 10
Rationale and objectives 11
Methodology and outline of the dissertation 12
CHAPTER 2 – A SCOPING REVIEW OF THE LITERATURE FEATURING RESEARCH ETHICS AND RESEARCH INTEGRITY CASES 15
Background 15
Methods 16
Results 19
Discussion 30
Conclusion 37
CHAPTER 3 – CHARACTERIZATION OF PHD STUDENTS AND DISCIPLINARY DIFFERENCES ON PERCEPTIONS AND ATTITUDE TOWARDS RCR 38
Background 38
Methods 39
Results 42
Discussion 57
Conclusion 62
CHAPTER 4 – PERCEPTIONS OF RESEARCH INTEGRITY CLIMATE IN HUNGARIAN UNIVERSITIES: RESULTS FROM A SURVEY AMONG ACADEMIC RESEARCHERS 63
Background 63
Methods 65
Results 67

Discussion ... 71
Conclusion ... 74
CHAPTER 5 – PERCEPTIONS OF PUBLICATION PRESSURE: SURVEY WITH RESEARCHERS FROM HUNGARIAN UNIVERSITIES ... 76
Background ... 76
Methods ... 77
Results ... 79
Discussion ... 82
Conclusion ... 84
CHAPTER 6 – DISSERTATION DISCUSSION ... 85
Main findings ... 85
Strengths and limitations ... 86
Suggestions for future research ... 87
Practical implications and recommendations ... 87
Concluding remarks ... 90
SUMMARY ... 91
REFERENCES ... 92
PUBLICATION LIST ... 102
KEYWORDS ... 104
ACKNOWLEDGMENTS ... 105
APPENDICES ... 106
Appendix 1 ... 106
Appendix 2 ... 110
Appendix 3 ... 111
Appendix 4 ... 112

CHAPTER 1 – INTRODUCTION AND DISSERTATION BACKGROUND

INTRODUCTION

The fundamental pillar of scientific progress has always been based on the reliability of the knowledge developed by others. Individual and collective adherence to the core principles such as reliability, honesty, respect, and accountability ensures the integrity of science and knowledge (ALLEA 2017). The commitment to these values by institutions and researchers represents responsible conduct of research (RCR) (Steneck 2006). Therefore, violating these fundamental values jeopardize the whole research enterprise.

Misconduct cases and questionable research practices (QRPs) have questioned the reliability and objectivity of science and undermined public trust in science (Armond et al. 2021). However, the responsibility to ensure the integrity of research goes beyond individual researchers. Research performing organizations (RPO), research funding organizations (RFO), scientific journals, sponsors, and all stakeholders in the research enterprise should encourage those values and maintain research environments to support integrity.

A growing body of evidence shows that QRPs and external aspects in the research environment, such as publication pressure, might be some of the causes of the lack of reproducible studies results (Baker 2016). Therefore, it is necessary to understand the research system, its players, and incentives to address research integrity in this complex context. For that, researchers and policymakers developed some frameworks to investigate, prevent, and correct them.

LITERATURE REVIEW

Concepts and definitions

In the 1980s and 1990s, widely publicized research misconduct cases shocked the scientific community (Armond et al. 2021). Until that time, little was known about research behavior, and misconduct cases were historically labeled as merely "bad apples". The “bad apple”

narrative views misconduct cases as an individual problem that affects few bad individuals and does not reflect the scientific community (Sovacool 2008). Accordingly, this narrative focused the policy efforts on the investigations and punishment of those involved in misconduct rather than preventive measures (Redman 2013). However, these cases motivated a growing body of research to further examine research practices, identify the causes and describe the consequences of research misbehaviors. Some important concepts were developed to classify those behaviors and the emerging field.

On one side of the research behavior line, is *responsible conduct of research* (RCR). RCR has been defined by Nicholas Steneck (2006) as “conducting research in ways that fulfill the professional responsibilities of researchers, as defined by their professional organizations, the institutions for which they work and, when relevant, the government and public”. To put it simply, “the ideal standard institutions and individuals endeavor to meet”. In this framework, he splits RCR into two subfields, research ethics (RE) and research integrity (RI). While the former addresses research behavior from the moral principles perspective, and the latter from the professional standards.

In the between, is the *Questionable research practices* (QRPs), also known as detrimental research practices. QRPs were defined as "actions that violate traditional values of the research enterprise, and that may be detrimental to the research process" (National Academies of Sciences and Medicine 2017). These practices have been pointed out as one of the main causes of the reproducibility crisis (Baker 2016), and they involve not only individual researchers but also institutions, sponsors, and journals. QRPs include selective reporting, bad authorship and citation practices, conflicts of interest, overinterpretation of study results, and p-hacking.

On the other side of the research behavior line, it would be the more deliberate non-compliant behavior, the *Research misconduct.* Research misconduct has been traditionally defined as Fabrication, Falsification, and Plagiarism, the so-called FFP (Steneck 2006). This definition applies when proposing, reviewing, reporting, or performing research. Misconduct cases are generally more easily detectable and receive significant attention from the media and academia (Armond et al. 2021). However, with new evidence, it

became clear that less serious and way more prevalent questionable research practices can be more damaging to scientific integrity (Martinson, Anderson, and de Vries 2005).

Consequences of research misbehavior

Quantifying the consequences of misconduct or QRPs can be difficult due to the variety of involved factors. The consequences can involve damage to the individuals, costs to the institutions or journals, financial costs, and social consequences. (National Academies of Sciences and Medicine 2017).

The damage to individuals includes those who engaged in misconduct and those who rely on the fraudulent article. Individuals involved in misconduct cases can face damages to their reputation and career and also direct implications such as paper retractions, fund returns, and degree revocation (Armond and Kakuk 2021). Moreover, institutions and journals might deal with reputational costs and prestige losses.

The financial costs include the grants supporting the research projects and the funds on research misconduct investigations. Although hard to estimate the direct costs of misconduct, a study on retracted papers (Stern et al. 2014) found that misconduct incurs high direct and indirect financial costs. These costs affect the funding organizations, institutions, and investigation committees. Although the financial costs are substantial, the study highlights that the social consequences can be even more detrimental.

Social consequences of misconduct can be severe, including loss of human life due to misinformation in the literature. A clear example is the infamous article by Wakefield et al. (1998) that mistakenly suggested a link between vaccination and autism. The emerging public debate on the vaccination of children has been strongly influenced by this article, which is responsible for low vaccination rates and preventable illnesses. Public trust in science is essential, and misconduct cases undermine it, posing a threat during a public health crisis, such as COVID-19 (Schwartz 2020, Sulik et al. 2021).

Furthermore, research misbehavior and QRPs might affect the reproducibility of study results, affecting research accuracy and reliability (Ioannidis 2005). For instance, effective policymaking should be built upon robust and reliable research to benefit society and

societal progress. Thus, policymaking based on flawed research results can lead to serious harm to society and jeopardize the effectiveness of public policies (Napoli and Karaganis 2010).

Causal factors of research misbehavior

The “bad apple” metaphor has traditionally protected science and perpetuated an ineffective way of dealing with misconduct by placing the burden only on individuals’ behavior (Redman 2013). Later this narrative moved to a broader understanding, including individual, organizational, and structural factors. (Davis, Riske-Morris, and Diaz 2007a, Sovacool 2008)

Due to the “bad apple” metaphor, individual factors were traditionally the most mentioned aspect involved in misconduct cases. These factors include mental and emotional problems, personality traits, and poor judgment. (Davis, Riske-Morris, and Diaz 2007b, Kornfeld 2012). For instance, an empirical study by Tijdink et al. (2016) found that Machiavellianism[1] is associated with self-reported research misbehavior. Narrative analyses by Davis, Riske-Morris, & Diaz (2007a) and Kornfeld (2012) have also highlighted narcissism, perfectionism, and sociopathy as risk factors for research misbehavior.

Organizational factors are certain working environment features that might undermine scientific integrity. These factors include the lack of clear guidelines and policies on RCR, poor mentoring, inadequate training, and organizational justice/injustice. *Organizational justice theory* is an important framework that addresses organizational factors (Martinson et al. 2010a). The theory associates individuals' perceptions about the "fairness" of decision-making and resource distribution within organizations and the behavioral consequences of those perceptions. This means that if people perceive injustice in their organization, they are more likely to behave in ways that, in their mind, compensates for the perceived unfairness. In a research climate where perceived injustice is high, researchers would be

[1] Machiavellianism can be defined as “a person’s tendency to be unemotional, detached from conventional morality and hence to deceive and manipulate others, to focus on unmitigated achievement and give high priority to own performances”, according to Christie R. and Geis F. (1970) Studies in Machiavellianism. NY: Academic Press.

expected to be more likely to engage in intentional research misconduct (falsification, fabrication, and plagiarism) or questionable research practices.

Structural factors represent a broader problem, extending beyond the individual or institutional factors and involving the research culture. A well-known structural factor is the “publish or perish” culture. Publication and competitive pressures have been appointed as strong influences to research misbehaviors (Fanelli 2010). However, this assumption is still controversial (Fanelli, Costas, and Lariviere 2015) The other structural factor is the scientific award system. However, this assumption is still controversial (Fanelli, Costas, and Lariviere 2015). The award system of science is another example of a structural factor. The quantitative approach to researcher assessment (e.g., using metrics) might also influence an individual’s behavior to play the system and cut corners (Sovacool 2008). This narrative highlights the shift of scientific values. While science was fundamentally based on values, such as trust, cooperation, reliability, discovery, and so on, the “commercialization of science” prioritizes competition, publication, and exploitation (Sovacool 2008). Sovacool highlights that this problem will persist as long as these phenomena exist.

Therefore, understanding the causes of research misbehavior is fundamental to address the matter and create solutions to foster research integrity.

Beyond research misconduct

Although research misconduct is the first thing that comes to mind when thinking about research integrity, research integrity is much more than misconduct. The research integrity field is more a matter of implementing mechanisms that improve research quality, relevance, and reliability. Incentives for open data, open access and data sharing, protocol registration, better record-keeping, better mentoring, bias minimization, and more rigorous experimental design are instances of mechanisms studied by the research field. The mechanisms rely on the evidence that bad research practices (e.g., p-hacking, low statistical power, publication bias, and poor record-keeping) are more prevalent than misconduct, and therefore, more detrimental to the integrity of science (Ioannidis 2005). Accordingly,

improvements to routine research practices are critical to achieving research integrity and robustness.

Science reliability and reproducibility have been tested and questioned (Ioannidis 2005, Baker 2016), and many times have found disturbing results. Although some can see the narrative of a “reproducibility crisis” as counterproductive (Fanelli 2018), several studies have pointed the problem and failed to replicate or reproduce scientific results. A recent reproducibility project tried to replicate high-impact cancer research papers, and half of the attempts failed despite their efforts (Mullard 2021). Although replication is extremely hard, the mounting evidence on reproducibility issues calls for new initiatives and tools to address them. An important point to consider is that even a completely honest researcher can commit some mistake and have self-deceptive behavior. Several examples in wrong data analysis (Nuzzo 2015, 2014) have shown how easy it is to fool ourselves and jump to conclusions when facing reasonable results. Understanding our own biases will help us develop tools to move towards a more robust science.

Fostering research integrity

Over the past years, several initiatives from various stakeholders emerged to foster research integrity. In Europe, dozens of research projects were funded by the European Commission to address research integrity and responsible research and innovation from different approaches.

This dissertation highlights here two different projects on which the results presented here are partially based. Both projects were funded by European Union’s Horizon 2020 and had the University of Debrecen as a partner.

The EnTIRE consortium is a collaboration between eleven partners. Its main aim was to create a platform to make the normative framework governing Research Ethics and Research Integrity accessible, the Embassy of Good Science (https://embassy.science) (Nature 2019). This platform development counted on the participation of different stakeholders, and it is now the official European platform for research ethics and integrity. The Embassy of Good Science is an interactive and self-sustainable platform in a wiki

format. However, the initial inputs and contents were also developed by the project. **Chapter 2** of this dissertation presents a scoping review of RERI cases that are available on the platform. The other project is the INTEGRITY project. The project had also involved eleven partners that aim to create a teaching methodology to empower students on RCR. This project involves different approaches, and **Chapter 3** of this dissertation shows some results from a survey developed to map students' knowledge and attitude on RCR to provide evidence-based input for the development of teaching tools.

There is no single approach to address RI and RCR. Only in Europe, there are dozens of networks and projects to foster the integrity of science. Accumulated evidence shows that tailored interventions to promote RCR, such as developing policies, guidelines, and education in RCR, should be the primary response to integrity concerns.

The recommendations to foster integrity involve the various stakeholders in the research enterprise. It includes creating and improving practices and policies to address research misbehaviors and monitor newly emerging threats to research integrity. Since RPOs have an essential role in promoting research integrity, they should identify their weaknesses and address them by maintaining high standards for research conduct and promoting a healthy work environment.

RATIONALE AND OBJECTIVES

Different initiatives can be employed to improve education and research practices. Yet, to ensure the effectiveness of these interventions, it is essential to develop evidence-based initiatives tailored to their needs and contexts. Many countries and institutions have acknowledged the importance of raising awareness of research integrity and developing procedures for dealing with its threats.

However, the public discourse and institutionalization of research integrity are still lagging in Hungary compared to other countries, and very little is known about research integrity in the country. Therefore, the main objective of this dissertation was to provide some evidence on research integrity and its differences across disciplines, most specifically to the Hungarian contexts, to drive future initiatives to foster research integrity in the country.

The objectives were the following:

- Map and characterize research ethics and integrity cases available in the literature.
- Characterize PhD students regarding understandings and perceptions towards RCR in Hungarian universities, and investigate disciplinary differences.
- Assess the prevalence of self-reported and observed misconduct and questionable research practices among PhD students in Hungary, and investigate disciplinary differences.
- Assess the academic researchers' perceptions of research integrity climate in Hungary and the differences across disciplines, and career stages.
- Assess the academic researchers' perceptions of publication pressure in Hungary and investigate the differences across disciplines, career stages, and gender.

METHODOLOGY AND OUTLINE OF THE DISSERTATION

Research integrity evidence can be investigated from different approaches and methodologies. For this dissertation, due to reasons discussed in the limitation section, we chose to use mainly a qualitative approach through survey questionnaires.

The dissertation initiates in **Chapter 2** by characterizing integrity and ethics cases. The aim was to explore how the literature presents research ethics and integrity cases and characterize and analyze them concerning the violations involved, sanctions, and field of science. For that, a scoping review on cases available in the academic literature was conducted. This chapter focuses primarily on individuals, providing information on prominent cases, violations, and sanctions. The results show that the literature consists mainly of FFP and prominent cases reports, mostly involving the biomedical sciences. But do the same characteristics apply to the Hungarian context? Is the prevalence of FFP higher than other misbehaviors? Are the biomedical sciences the most likely to engage in misbehavior? To address these questions, **Chapter 3** studies the Hungarian context. The aim was to characterize the PhD students in Hungary concerning perceptions and attitudes towards RCR and assess potential disciplinary differences. A survey with 200 PhD students from different universities in Hungary was conducted. The questionnaire was developed as

part of the INTEGRITY project to map comprehension of rules in RCR, self-perceived knowledge, education on RCR, self-reported and observed behavior towards RCR. This questionnaire tried not only to assess misconduct but also questionable research practices and grey zones. It is reasoned that these aspects could differ across disciplines, and they should be investigated to develop initiatives to foster research integrity. In general, **Chapters 2** and **3** places the focus on individual researchers, somehow reinforcing the "bad apple" narrative. However, structural and institutional factors play a significant role in influencing someone's misbehavior (Martinson et al. 2010a). Therefore, in the following chapters, structural and institutional factors were explored.

Chapter 4 assesses how scientists experience the research integrity climate, stratified for academic rank and scientific field, in three large Hungarian universities with doctoral schools covering the full spectrum of disciplines. The integrity climate was assessed using the SOURCE© (Survey of Organizational Research Climate) (Martinson, Thrush, and Lauren Crain 2013). This survey was developed based on the theoretical framework of organizational justice (Martinson et al. 2010a). The survey consists of seven scales, and it reflects factors such as integrity policy, resources, communication, and mentoring. The findings show that postdocs and assistant professors perceived the integrity climate more negatively on every scale. Moreover, the results indicate that responsible research conduct in Hungary is impacted by non-ideal research conditions related to resources, pressure (pressure to publish and obtain funding), and interpersonal relations.

In light of these results, it was possible to see that other external factors, such as publication pressure, could be relevant to the Hungarian context. The reward systems in academia and the way science is disseminated are significant contributors to the "publish or perish" culture. Excessive pressure to publish can lead to research flaws and spins and might also affect the likelihood of QRPs. Therefore, **Chapter 5** assessed perceived publication pressure and its relationship with academic rank, gender, and scientific field in Hungary. The revised Publication Pressure Questionnaire (PPQr) (Haven, de Goede, et al. 2019) was used to measure the perceived publication pressure. The questionnaire consists of three scales, stress, attitude, and resources.

Based on the findings, one can see that disciplines, ranks, and gender are affected differently by individual, structural, and institutional factors. **Chapter 6** discusses what these results mean to the Hungarian contexts. What institutions and researchers can do to foster integrity? What are the current practices for RCR?

CHAPTER 2 – A SCOPING REVIEW OF THE LITERATURE FEATURING RESEARCH ETHICS AND RESEARCH INTEGRITY CASES [2]

BACKGROUND

There has been an increase in academic interest in research ethics (RE) and research integrity (RI) over the past decade. This is due, among other reasons, to the changing research environment with new and complex technologies, increased pressure to publish, greater competition in grant applications, increased university-industry collaborative programs, and growth in international collaborations (National Academies of Sciences and Medicine 2017). In addition, part of the academic interest in RE and RI is due to highly publicized cases of misconduct (Davis, Riske-Morris, and Diaz 2007a).

There is a growing body of published RE and RI cases, which may contribute to public attitudes regarding both science and scientists (Ampollini and Bucchi 2020). Different approaches have been used in order to analyse RE and RI cases. Studies focusing on ORI files (Office of Research Integrity) (Davis, Riske-Morris, and Diaz 2007a), retracted papers (Hesselmann et al. 2017), quantitative surveys (Martinson, Anderson, and de Vries 2005), data audits (Loikith and Bauchwitz 2016), and media coverage (Ampollini and Bucchi 2020) have been conducted to understand the context, causes and consequences of these cases.

Analyses of RE and RI cases often influence policies on responsible conduct of research (National Academies of Sciences and Medicine 2017). Moreover, details about cases facilitate a broader understanding of issues related to RE and RI and can drive interventions to address them. Currently, there are no comprehensive studies that have collected and evaluated the RE and RI cases available in the academic literature. This review has been

[2] This chapter was published at BMC Medical Ethics. Armond, A. C. V., Gordijn, B., Lewis, J., Hosseini, M., Bodnár, J. K., Holm, S., & Kakuk, P. (2021). A scoping review of the literature featuring research ethics and research integrity cases. *BMC Med Ethics, 22*(1), 50. doi:10.1186/s12910-021-00620-8

developed by members of the EnTIRE consortium[3] to generate information on the cases that will be made available on the Embassy of Good Science platform (www.embassy.science). Two separate analyses have been conducted. The first analysis uses identified research articles to explore how the literature presents cases of RE and RI, in relation to the year of publication, country, article genre, and violation involved. The second analysis uses the cases extracted from the literature in order to characterize the cases and analyze them concerning the violations involved, sanctions, and field of science.

METHODS

This scoping review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and PRISMA Extension for Scoping Reviews (PRISMA-ScR). The full protocol was pre-registered and it is available at https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5bde92120&appId=PPGMS.

Eligibility

Articles with case(s) involving a violation of, or misbehavior, poor judgment, or detrimental research practice in relation to a normative framework, were included. Cases unrelated to scientific activities, research institutions, academic or industrial research and publication were excluded. Articles that did not contain a substantial description of the case were also excluded.

A normative framework consists of explicit rules, formulated in laws, regulations, codes, and guidelines, as well as implicit rules, which structure local research practices and influence the application of explicitly formulated rules. Therefore, if a case involves a violation of, or misbehavior, poor judgment or detrimental research practice in relation to a normative framework, then it does so on the basis of explicit and/or implicit rules governing RE and RI practice.

[3] EnTIRE project (Mapping Normative Frameworks for Ethics and Integrity of Research) has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement N741782. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Search strategy

A search was conducted in PubMed, Web of Science, SCOPUS, JSTOR, Ovid, and Science Direct in March 2018, without any language or date restrictions. Two parallel searches were performed with two sets of medical subject heading (MeSH) terms, one for RE and another for RI. The parallel searches generated two sets of data thereby enabling us to analyse and further investigate the overlaps in, differences in, and evolution of, the representation of RE and RI cases in the academic literature. The terms used in the first search were: (("research ethics") AND (violation OR unethical OR misconduct)). The terms used in the parallel search were: (("research integrity") AND (violation OR unethical OR misconduct)). The search strategy's validity was tested in a pilot search, in which different keyword combinations and search strings were used, and the abstracts of the first hundred hits in each database were read.

After searching the databases with these two search strings, the titles and abstracts of extracted items were read by three contributors independently (ACVA, PK, and KB). Articles that could potentially meet the inclusion criteria were identified. After independent reading, the three contributors compared their results to determine which studies were to be included in the next stage. In case of a disagreement, items were reassessed in order to reach a consensus. Subsequently, qualified items were read in full.

Data extraction

Data extraction processes were divided by three assessors (ACVA, PK and KB). Each list of extracted data generated by one assessor was cross-checked by the other two. In case of any inconsistencies, the case was reassessed to reach a consensus. The following categories were employed to analyse the data of each extracted item (where available): (I) author(s); (II) title; (III) year of publication; (IV) country (according to the first author's affiliation); (V) article genre; (VI) year of the case; (VII) country in which the case took place; (VIII) institution(s) and person(s) involved; (IX) field of science (FoS-OECD classification) (OECD 2007); (X) types of violation (see below); (XI) case description; and (XII) consequences for persons or institutions involved in the case.

Two sets of data were created after the data extraction process. One set was used for the analysis of articles and their representation in the literature, and the other set was created for the analysis of cases. In the set for the analysis of articles, all eligible items, including duplicate cases (cases found in more than one paper, e.g., Hwang case, Baltimore case) and fictional cases were included. The aim was to understand the historical aspects of violations reported in the literature as well as the paper genre in which cases are described and discussed. For this set, the variables of the year of publication (III); country (IV); article genre (V); and types of violation (X) were analysed.

For the analysis of cases, all duplicated cases, fictitious cases, and cases that did not contain enough information about particularities to differentiate them from others (e.g., names of the people or institutions involved, country, date) were excluded. In this set, prominent cases (i.e., those found in more than one paper) were listed only once, generating a set containing solely unique cases. These additional exclusion criteria were applied to avoid multiple representations of cases. For the analysis of cases, the variables: (VI) year of the case; (VII) country in which the case took place; (VIII) institution(s) and person(s) involved; (IX) field of science (FoS-OECD classification); (X) types of violation; (XI) case details; and (XII) consequences for persons or institutions involved in the case were considered.

Article genre classification

We used ten categories to capture the differences in genre. In most cases, the articles were classified following the classification provided by the journal. For instance, we included a case description in a “news” genre if a case was published in the news section of a scientific journal or newspaper. Although we have not developed a search strategy for newspaper articles, some of them (e.g., New York Times) are indexed in scientific databases such as Pubmed. The same method was used to allocate case descriptions to “editorial”, “commentary”, “misconduct notice”, “retraction notice”, “review”, “letter” or “book review”. We applied the “case analysis” genre if a case description included a normative analysis of the case. The “educational” genre was used when a case description was incorporated to illustrate RE and RI guidelines or institutional policies.

Categorization of violations

For the extraction process, we used the articles' own terminology when describing violations/ethical issues involved in the event (e.g., plagiarism, falsification, ghost authorship, conflict of interest, etc.) to tag each article. In case the terminology was incompatible with the case description, other categories were added to the original terminology for the same case. Subsequently, the resulting list of terms was standardized using the list of major and minor misbehaviors developed by Bouter and colleagues (Bouter et al. 2016). This list consists of 60 items classified into four categories: Study design, data collection, reporting, and collaboration issues (Appendix 1).

RESULTS

Systematic search

A total of 11,641 records were identified through the RE search and 3,078 in the RI search. The results of the parallel searches were combined, and the duplicates removed. The remaining 10,556 records were screened, and at this stage, 9,750 items were excluded because they did not fulfil the inclusion criteria. 806 items were selected for full-text reading. Subsequently, 388 articles were included in the qualitative synthesis (Figure 1)[4].

---

[4] The dataset supporting the conclusions of this article is available in the Open Science Framework (OSF) repository in https://osf.io/3xatj/?view_only=313a0477ab554b7489ee52d3046398b9.

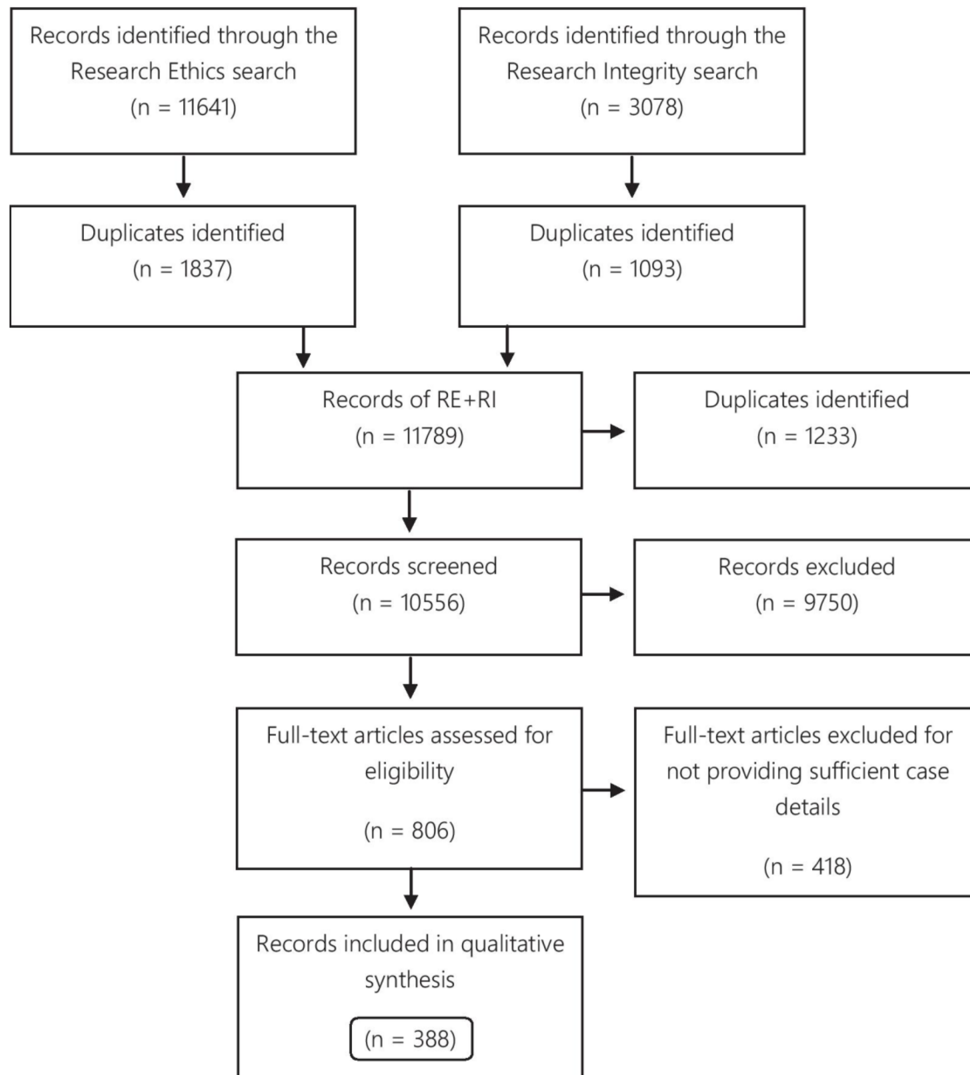


Figure 1. Flow diagram showing the article selection process

Of the 388 articles, 157 were only identified via the RE search, 87 exclusively via the RI search, and 144 were identified via both search strategies. The eligible articles contained 500 case descriptions, which were used for the analysis of the publications articles analysis. 256 case descriptions discussed the same 50 cases. The Hwang case was the most frequently described case, discussed in 27 articles. Furthermore, the top 10 most described cases were found in 132 articles (Table 1).

Table 1. Top 10 most described cases.

Cases	Articles	Date range
1.Hwang	27	2005-2016
2.Baltimore/Imanishi-kari	24	1990-2007
3.Gallo	21	1990-2010
4.Fisher/Poisson	12	1994-1997
5. Schön	10	2002-2014
6.Luk Van Parijs	9	1998-2011
7.Poehlman	8	2005-2010
8.Boldt	8	2011-2014
9.Wakefield	7	2004-2013
10.CNEP	6	2006-2010

For the analysis of cases, 206 (41.2% of the case descriptions) duplicates were excluded, 56 (11.2%) cases were excluded for not providing enough information to distinguish them from other cases.

1. Violations

1.1 Analysis of the articles

The categories used to classify the violations include those that pertain to the different kinds of scientific misconduct (falsification, fabrication, plagiarism), detrimental research practices (authorship issues, duplication, peer-review, errors in experimental design, and mentoring), and "other misconduct" (according to the definitions from the National Academies of Sciences and Medicine) (National Academies of Sciences and Medicine 2017). Each case could involve more than one type of violation. The majority of cases presented more than one violation or ethical issue, with a mean of 1.56 violations per case. Figure 2 presents the frequency of each violation tagged to the articles. Falsification and fabrication were the most frequently tagged violations. The violations accounted

respectively for 29.1% and 30.0% of the number of taggings (n = 780), and they were involved in 46.8% and 45.4% of the articles (n = 500 case descriptions). Problems with informed consent represented 9.1% of the number of taggings and 14% of the articles, followed by patient safety (6.7% and 10.4%) and plagiarism (5.4% and 8.4%). Detrimental research practices, such as authorship issues, duplication, peer-review, errors in experimental design, mentoring, and self-citation were mentioned cumulatively in 7.0% of the articles.

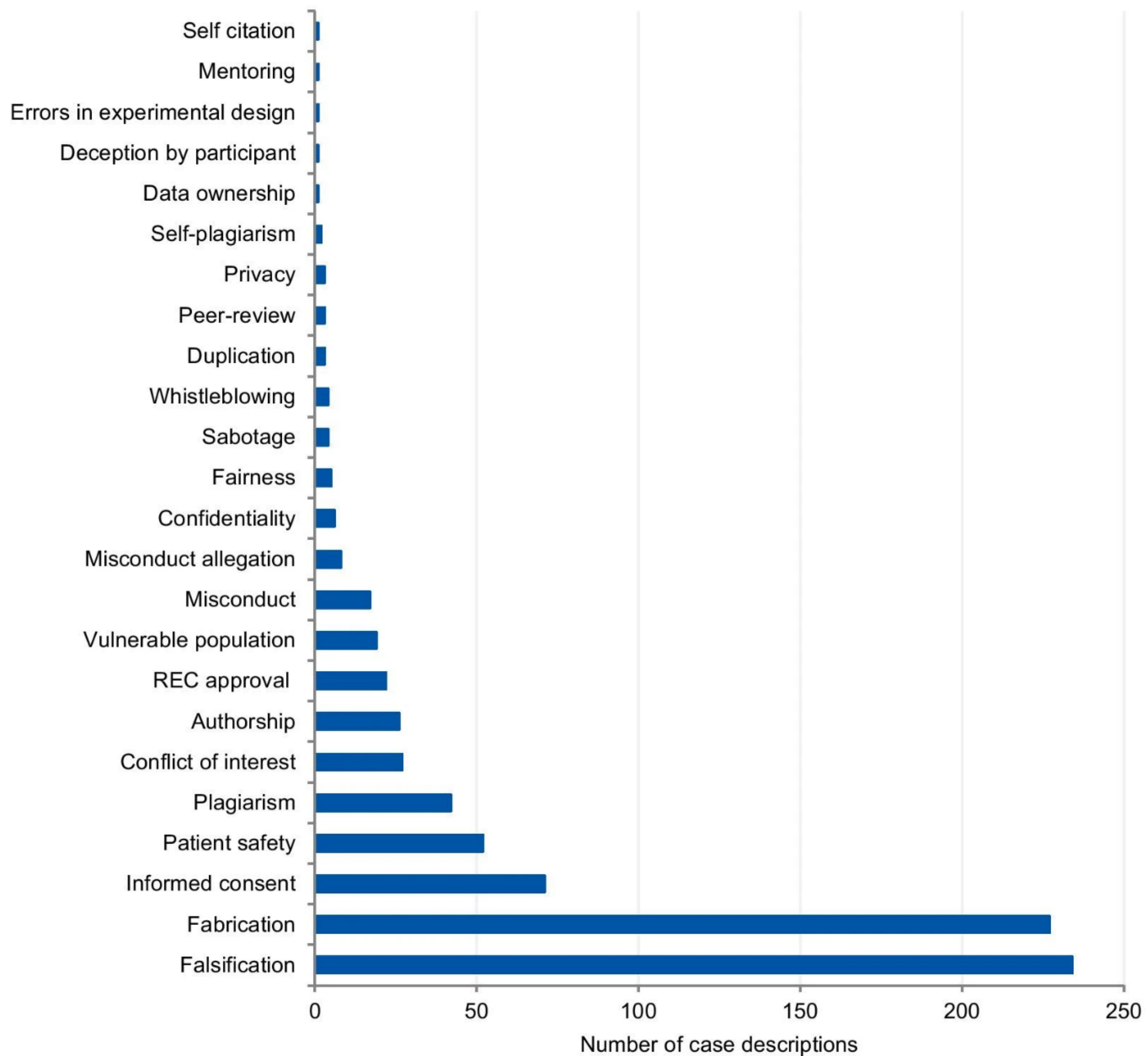


Figure 2. Percentage of tagged violations from the article analysis

1.2 Analysis of the cases

Figure 3 presents the frequency and percentage of each violation found in the cases. Each case could include more than one item from the list. The 238 cases were tagged 305 times, with a mean of 1.28 items per case. Fabrication and falsification were the most frequently tagged violations (44.9%), involved in 57.7% of the cases (n = 238). The non-adherence to pertinent laws and regulations, such as lack of informed consent and REC approval, was the second most frequently tagged violation (15.7%) and involved in 20.2% of the cases. Patient safety issues were the third most frequently tagged violations (11.1%), involved in 14.3% of the cases, followed by plagiarism (6.9% and 8.8%). The list of major and minor misbehaviors (Bouter et al. 2016) classifies the items into study design, data collection, reporting, and collaboration issues. Our results show that 56.0% of the tagged violations involved issues in reporting, 16.4% in data collection, 15.1% involved collaboration issues, and 12.5% in the study design. The items in the original list that were not listed in the results were not involved in any case collected.

Major and minor misbehavior items	Number of cases	Percentage	Percentage of cases (n=238)	Domain
Selectively delete data. modify data or add fabricated data after performing initial data-analyses	137	44.92	57.56	Reporting
Not adhere to pertinent laws and regulations	48	15.74	20.17	Data collection
Ignore substantial safety risks of the study to participants. workers or environment	34	11.15	14.29	Study design
Use published ideas or phrases of others without referencing	21	6.89	8.82	Collaboration
Modify the results or conclusions of a study due to pressure of a sponsor	9	2.95	3.78	Reporting
Failure to disclose a relevant financial or intellectual conflict of interest	5	1.64	2.10	Reporting
Handle existing conflicts of interest inadequately	5	1.64	2.10	Reporting
Refuse to respond to an allegation of a breach of research integrity	5	1.64	2.10	Collaboration
Not report all study protocol-stipulated results	4	1.31	1.68	Reporting
Present grossly misleading information in a grant application	3	0.98	1.26	Study design
Duplicate publication without disclosure	3	0.98	1.26	Reporting
Failure to disclose a sponsor of the study	3	0.98	1.26	Reporting
Omit a contributor who deserves authorship	3	0.98	1.26	Collaboration
Turn a blind eye to putative breaches of research integrity by others	2	0.66	0.84	Collaboration
Re-use parts of your own publications without referencing	2	0.66	0.84	Collaboration
Unfairly review papers. grant applications or colleagues applying for promotion	2	0.66	0.84	Collaboration
Add an author who doesn't qualify for authorship	2	0.66	0.84	Collaboration
Demand or accept an authorship for which you don't qualify	2	0.66	0.84	Collaboration
Submit or resubmit a paper or grant application without consent from all authors	2	0.66	0.84	Collaboration
Choose a clearly inadequate research design or using evidently unsuitable measurement instruments	1	0.33	0.42	Study design
Inadequately handle or store data or (bio)materials	1	0.33	0.42	Data collection
Keep inadequate notes of the research process	1	0.33	0.42	Data collection
Report an unexpected finding as having been hypothesized from the start	1	0.33	0.42	Reporting
Not report clearly relevant details of study methods	1	0.33	0.42	Reporting
Not report replication problems	1	0.33	0.42	Reporting
Selectively cite or cite your own work to improve citation metrics	1	0.33	0.42	Reporting
Re-use of previously published data without disclosure	1	0.33	0.42	Reporting
Refuse to share data with bona fide colleagues	1	0.33	0.42	Collaboration
Use unpublished ideas or phrases of others without their permission	1	0.33	0.42	Collaboration
Review your own papers	1	0.33	0.42	Collaboration
Insufficiently supervise or mentor junior coworkers	1	0.33	0.42	Collaboration
Be grossly unfair to your collaborators	1	0.33	0.42	Collaboration

Figure 3. Frequency and percentage of major and minor misbehavior items from the analysis of cases

2. Article genre

The articles were mostly classified into “news” (33.0%), followed by “case analysis” (20.9%), “editorial” (12.1%), “commentary” (10.8%), “misconduct notice” (10.3%), “retraction notice” (6.4%), “letter” (3.6%), “educational paper” (1.3%), “review” (1%), and “book review” (0.3%) (Figure 4). The articles classified into “news” and “case analysis” included predominantly prominent cases. Items classified into “news” often explored all the investigation findings step by step for the associated cases as the case progressed through investigations, and this might explain its high prevalence. The case analyses included mainly normative assessments of prominent cases. The misconduct and retraction notices included the largest number of unique cases, although a relatively large portion of the retraction and misconduct records could not be included because of insufficient case details. The articles classified into “editorial”, “commentary” and “letter” also included unique cases.

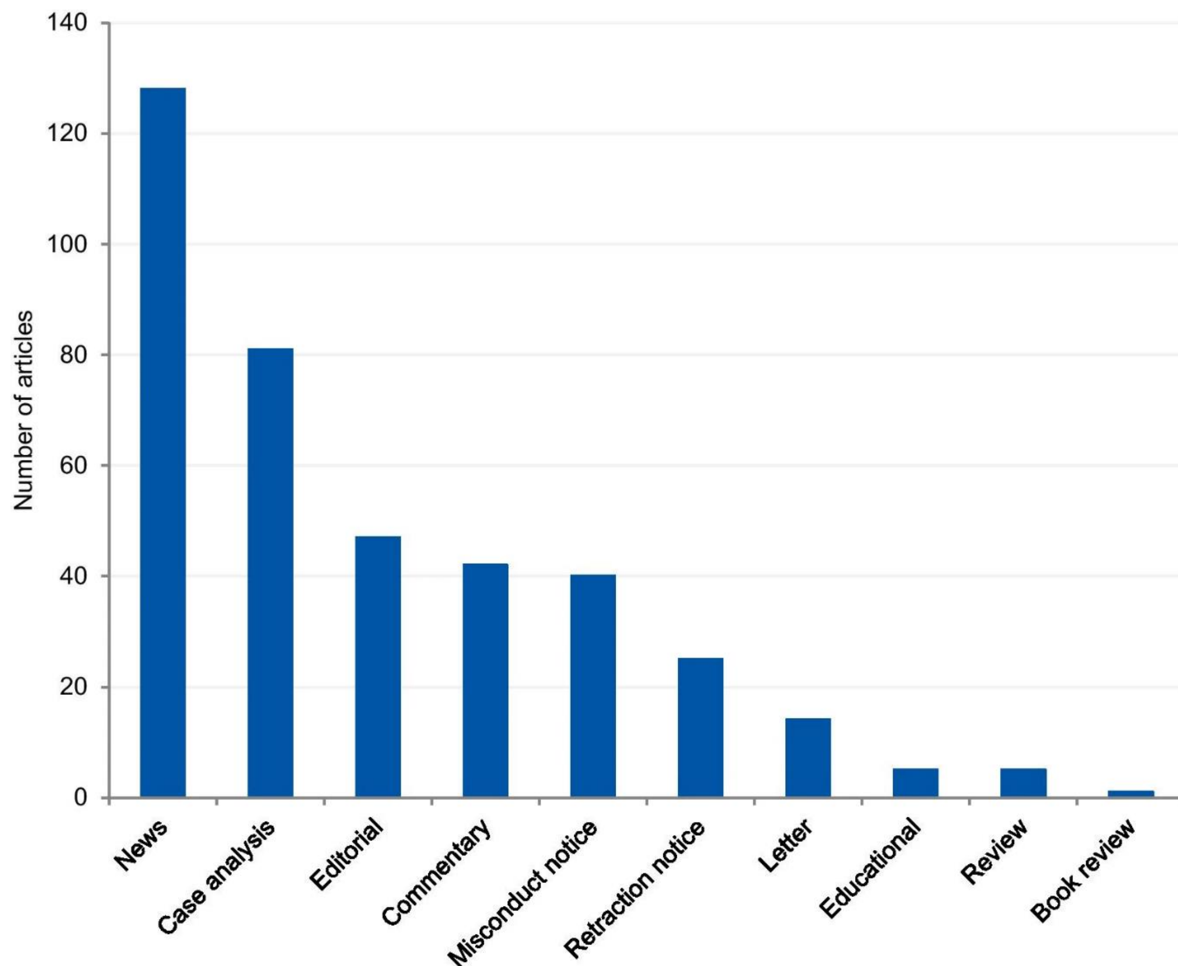


Figure 4. Frequency of article genre

3. Date

3.1 Article analysis

The dates of the eligible articles range from 1983 to 2018 with notable peaks between 1990 and 1996, most probably associated with the Gallo (Greenberg 1995) and Imanishi-Kari cases (Dresser 1997), and around 2005 with the Hwang (Hong 2016), Wakefield (Opel, Diekema, and Marcuse 2011), and CNEP trial cases (Wells 2010) (Figure 5). The trend line shows an increase in the number of articles over the years.

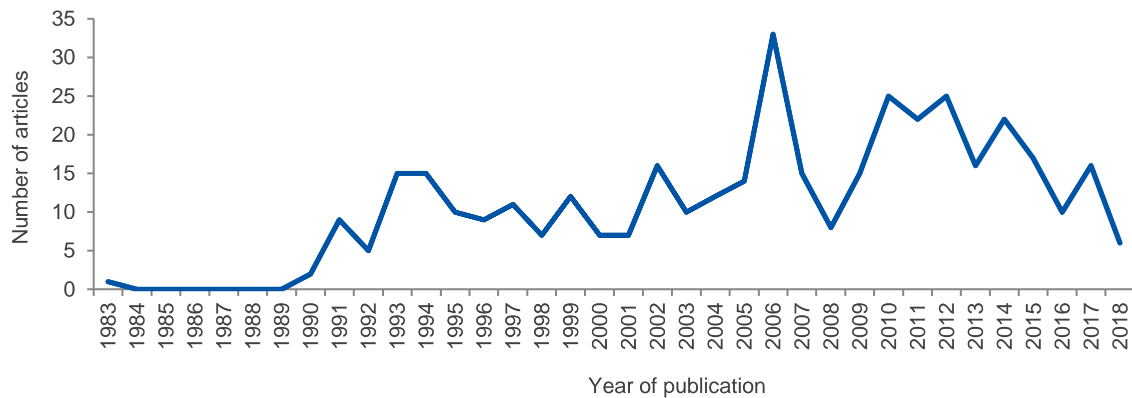


Figure 5. Frequency of articles according to the year of publication

3.2 Case analysis

The dates of included cases range from 1798 to 2016. Two cases occurred before 1910, one in 1798 and the other in 1845. Figure 6 shows the number of cases per year from 1910. An increase in the curve started in the early 1980s, reaching the highest frequency in 2004 with 13 cases.

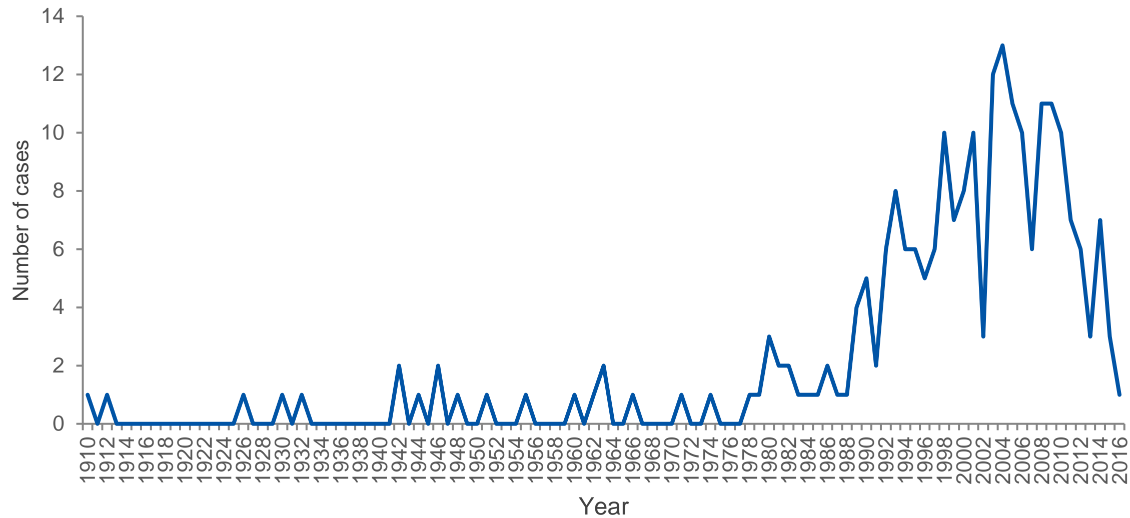


Figure 6. Frequency of cases per year

4. Geographical distribution

4.1 Article analysis

The first analysis concerned the authors' affiliation and the corresponding author's address. Where the article contained more than one country in the affiliation list, only the first author's location was considered. Eighty-one articles were excluded because the authors' affiliations were not available, and 307 articles were included in the analysis. The articles originated from 26 different countries (Appendix 2). Most of the articles emanated from the USA and the UK (61.9% and 14.3% of articles, respectively), followed by Canada (4.9%), Australia (3.3%), China (1.6%), Japan (1.6%), Korea (1.3%), and New Zealand (1.3%). Some of the most discussed cases occurred in the USA; the Imanishi-Kari, Gallo, and Schön cases (Dresser 1997, Greenberg 1995). Intensely discussed cases are also associated with Canada (Fisher/Poisson and Olivieri cases), the UK (Wakefield and CNEP trial cases), South Korea (Hwang case), and Japan (RIKEN case) (Opel, Diekema, and Marcuse 2011, Normile 2014). In terms of percentages, North America and Europe stand out in the number of articles (Figure 7).

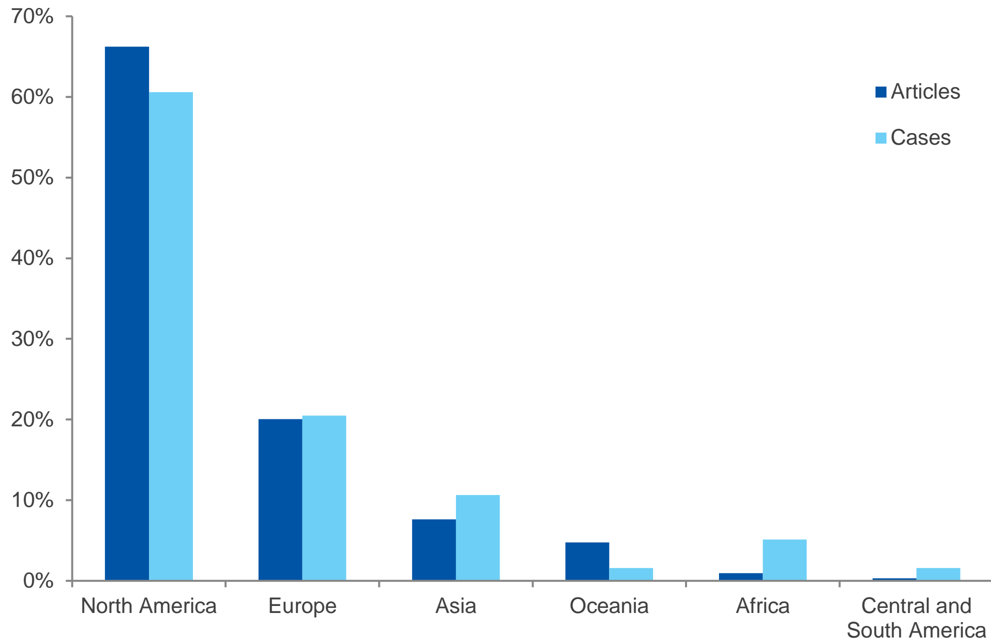


Figure 7. Percentage of articles and cases by continent

4.2 Case analysis

The case analysis involved the location where the case took place, taking into account the institutions involved in the case. For cases involving more than one country, all the countries were considered. Three cases were excluded from the analysis due to insufficient information. In the case analysis, 40 countries were involved in 235 different cases (Appendix 3). Our findings show that most of the reported cases occurred in the USA and the United Kingdom (59.6% and 9.8% of cases, respectively). In addition, a number of cases occurred in Canada (6.0%), Japan (5.5%), China (2.1%), and Germany (2.1%). In terms of percentages, North America and Europe stand out in the number of cases (Figure 7). To enable comparison, we have additionally collected the number of published documents according to country distribution, available on SCImago Journal & Country Rank (SCImago). The numbers correspond to the documents published from 1996 to 2019. The USA occupies the first place in the number of documents, with 21.9%, followed by China (11.1%), UK (6.3%), Germany (5.5%), and Japan (4.9%).

5. Field of science

The cases were classified according to the field of science. Four cases (1.7%) could not be classified due to insufficient information. Where information was available, 80.8% of cases were from the Medical and Health Sciences, 11.5% from the Natural Sciences, 4.3% from Social Sciences, 2.1% from Engineering and Technology, and 1.3% from Humanities (Figure 8). Additionally, we have retrieved the number of published documents according to scientific field distribution, available on SCImago (SCImago). Of the total number of scientific publications, 41.5% are related to natural sciences, 22% to engineering, 25.1% to health and medical sciences, 7.8% to social sciences, 1.9% to agricultural sciences, and 1.7 % to the humanities.

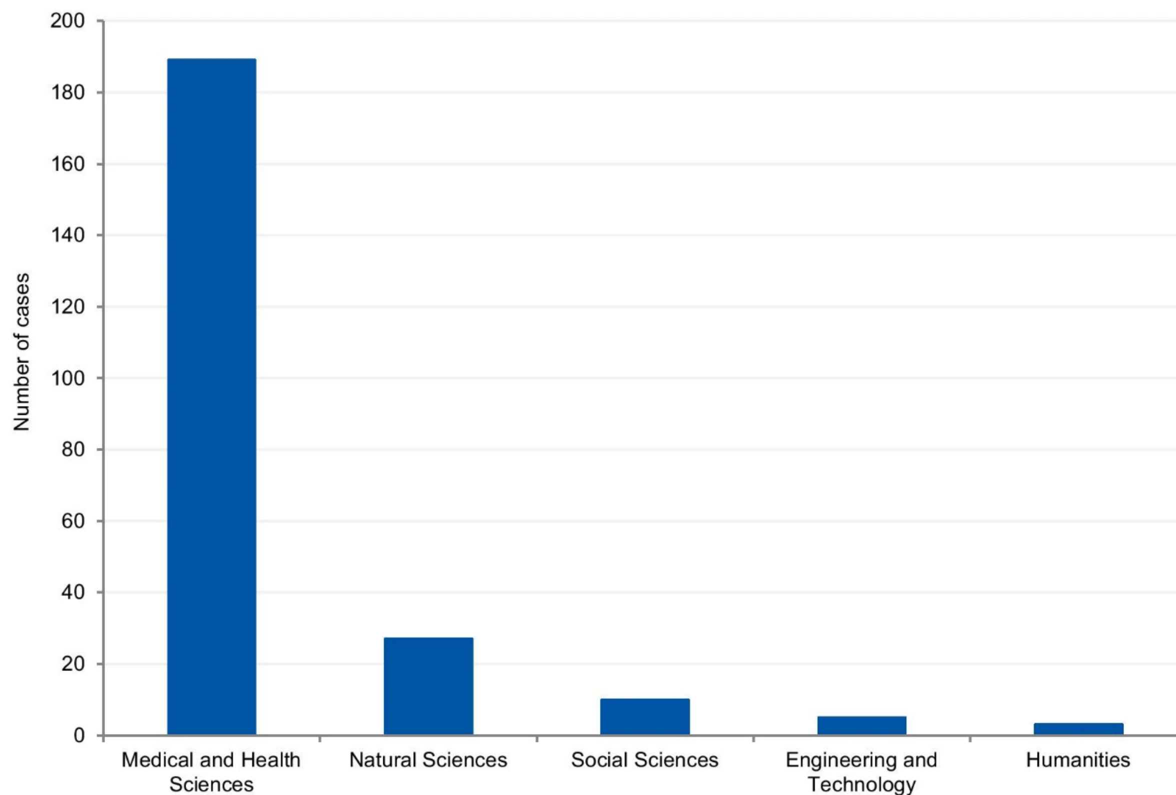


Figure 8. Percentage of cases by field of science

6. Sanctions

This variable aimed to collect information on possible consequences and sanctions imposed by funding agencies, scientific journals and/or institutions. 97 cases could not be classified due to insufficient information. 141 cases were included. Each case could potentially include more than one outcome. Paper retraction was the most prevalent sanction (45.4%), followed by exclusion from funding applications (35.5%) (Table 2).

Table 2. Frequency of different sanctions in the analysis of cases.

Sanction	Frequency	Percentage (%)	Percentage of cases (n=141) (%)
Paper retraction	64	33.3	45.4
Excluded from fund applications	50	26.0	35.5
Barred from service	22	11.5	15.6
Fired or suspended	18	9.4	12.8
Paper correction	12	6.3	8.5
Resignation	7	3.6	5.0
Trial	6	3.1	4.3
Manuscript rejection	5	2.6	3.5
Prison	3	1.6	2.1
Study halted	3	1.6	2.1
Fines / Restitution	2	1.0	1.4

DISCUSSION

RE and RI cases have been increasingly discussed publicly, affecting public attitudes towards scientists and raising awareness about ethical issues, violations, and their wider consequences (Martinson, Anderson, and de Vries 2005). Different approaches have been applied in order to quantify and address research misbehaviors (Martinson, Anderson, and de Vries 2005, Fanelli 2009, Steneck 2006, DuBois et al. 2013). However, most cases are investigated confidentially and the findings remain undisclosed even after the investigation (National Academy of Sciences, Institute of Medicine Panel on Scientific, and the Conduct of 1993) (DuBois et al. 2013, National Academy of Sciences, Institute of Medicine Panel on Scientific, and the Conduct of Research 1992). Therefore, the study aimed to collect the

RE and RI cases available in the scientific literature, understand how the cases are discussed, and identify the potential of case descriptions to raise awareness on RE and RI.

Articles

We collected and analyzed 500 detailed case descriptions from 388 articles and our results show that they mostly relate to extensively discussed and notorious cases. Nearly half of all included cases were mentioned in at least two different articles and the top-ten most commonly mentioned cases were discussed in 132 articles.

The prominence of certain cases in the literature, based on the number of duplicated cases we found (e.g., Hwang case), can be explained by the type of article in which cases are discussed and the type of violation involved in the case. In the article genre analysis, 33% of the cases were described in the news section of scientific publications. Our findings show that almost all article genres discuss those cases that are new and in vogue. Once the case appears in the public domain, it is intensely discussed in the media and by scientists, and some prominent cases have been discussed for more than 20 years (Table 1). Misconduct and retraction notices were exceptions in the article genre analysis, as they presented mostly unique cases. The misconduct notices were mainly found on the NIH repository, which is indexed in the searched databases. Some federal funding agencies like NIH usually publicize investigation findings associated with the research they fund. The results derived from the NIH repository also explains the large proportion of articles from the US (61.9%). However, in some cases, only a few details are provided about the case. For cases that have not received federal funding and have not been reported to federal authorities, the investigation is conducted by local institutions. In such instances, the reporting of findings depend on each institution's policy and willingness to disclose information (Bauchner et al. 2018). The other exception involves retraction notices. Despite the existence of ethical guidelines (COPE Council 2019), there is no uniform and common approach to how a journal should report a retraction. The Retraction Watch website suggests two lists of information that should be included in a retraction notice to satisfy the minimum and optimum requirements (COPE Council 2019, Retraction Watch 2015). As well as disclosing the reason for the retraction and information regarding the

retraction process, optimal notices should include: (I) the date when the journal was first alerted to potential problems; (II) details regarding institutional investigations and associated outcomes; (III) the effects on other papers published by the same authors; (IV) statements about more recent replications only if and when these have been validated by a third party; (V) details regarding the journal's sanctions; and (VI) details regarding any lawsuits that have been filed regarding the case. The lack of transparency and information in retraction notices was also noted in studies that collected and evaluated retractions (Fang, Steen, and Casadevall 2012). According to Resnik and Dinse (2013), retractions notices related to cases of misconduct tend to avoid naming the specific violation involved in the case. This study found that only 32.8% of the notices identify the actual problem, such as fabrication, falsification, and plagiarism, and 58.8% reported the case as replication failure, loss of data, or error. Potential explanations for euphemisms and vague claims in retraction notices authored by editors could pertain to the possibility of legal actions from the authors, honest or self-reported errors and lack of resources to conduct thorough investigations. In addition, the lack of transparency can also be explained by the conflicts of interests of the article's author(s), since the notices are often written by the authors of the retracted article.

The analysis of violations/ethical issues shows the dominance of fabrication and falsification cases (58.8%), and also explains the high prevalence of prominent cases. Non-adherence to laws and regulations, such as REC approval, informed consent, and data protection, was the second most prevalent issue (13%), followed by patient safety (6.7%), plagiarism (5.9%), and conflicts of interest (3.5%). The prevalence of the five most common violations in the case analysis was higher than the prevalence found in the analysis of articles. The only exceptions are fabrication and falsification cases, which made up 45% of the cases and were presented in 59.1% of the articles. The prevalence of violations involving non-adherence to laws and regulations, patient safety, plagiarism, and conflicts of interests was disproportionately higher in the analysis of cases compared to the analysis of articles that involved the same violations. This disproportion shows a predilection for the publication of discussions related to fabrication and falsification when compared to other serious violations. Complex cases involving these types of violations make good headlines and this follows a custom pattern of writing about cases that catch the public and media's attention (de Vries, Anderson, and Martinson 2006). The way cases of RE and RI violations

are explored in the literature gives a sense that only a few scientists are "the bad apples" and they are usually discovered, investigated, and sanctioned accordingly. This implies that the integrity of science, in general, remains relatively untouched by these violations. However, studies on misconduct determinants show that scientific misconduct is a systemic problem, which involves not only individual factors, but structural and institutional factors as well, and that a combined effort is necessary to change this scenario (Sovacool 2008, Haven, Tijdink, et al. 2019).

Analysis of cases

Date

A notable increase in RE and RI cases occurred in the 1990s, with a gradual increase until approximately 2006. This result is in agreement with studies that evaluated paper retractions (Trikalinos, Evangelou, and Ioannidis 2008, Fang, Steen, and Casadevall 2012). Although our study did not focus only on retractions, the trend is similar. This increase in cases should not be attributed only to the increase in the number of publications, since studies that evaluated retractions show that the percentage of retraction due to fraud has increased almost ten times since 1975, compared to the total number of articles. Our results also show a gradual reduction in the number of cases from 2011 and a greater drop in 2015. However, this reduction should be considered cautiously because many investigations take years to complete and have their findings disclosed. ORI has shown that from 2001 to 2010 the investigation of their cases took an average of 20.48 months with a maximum investigation time of more than 9 years (Fang, Steen, and Casadevall 2012).

Geographical distribution

The countries from which most cases were reported were the USA (59.6%), the UK (9.8%), Canada (6.0%), Japan (5.5%), and China (2.1%). When analyzed by continent, the highest percentage of cases took place in North America, followed by Europe, Asia, Oceania, Latin America, and Africa. The predominance of cases from the USA is predictable, since the country publishes more scientific articles than any other country, with 21.8% of the total documents, according to SCImago (SCImago). However, the same interpretation does not

apply to China, which occupies the second position in the ranking, with 11.2%. These differences in the geographical distribution were also found in a study that collected published research on research integrity (Aubert Bonn and Pinxten 2019). The results found by Bonn and Pinxten (2019) show that studies in the United States accounted for more than half of the sample collected, and although China is one of the leaders in scientific publications, it represented only 0.7% of the sample. Our findings can also be explained by the search strategy that included only keywords in English. Since the majority of RE and RI cases are investigated and have their findings locally disclosed, the employment of English keywords and terms in the search strategy is a limitation. Moreover, our findings do not allow us to draw inferences regarding the incidence or prevalence of misconduct around the world. Instead, it shows where there is a culture of publicly disclosing information and openly discussing RE and RI cases in English documents.

Scientific field analysis

The results show that 80.8% of reported cases occurred in the medical and health sciences whilst only 1.3% occurred in the humanities. This disciplinary difference has also been observed in studies on research integrity climates. A study conducted by Haven and colleagues (Haven, Tijdink, et al. 2019), associated seven subscales of research climate with the disciplinary field. The subscales included: (1) Responsible Conduct of Research (RCR) resources, (2) regulatory quality, (3) integrity norms, (4) integrity socialization, (5) supervisor/supervisee relations, (6) (lack of) integrity inhibitors, and (7) expectations. The results, based on the seven subscale scores, show that researchers from the humanities and social sciences have the lowest perception of the RI climate. By contrast, the natural sciences expressed the highest perception of the RI climate, followed by the biomedical sciences. There are also significant differences in the depth and extent of the regulatory environments of different disciplines (e.g., the existence of laws, codes of conduct, policies, relevant ethics committees, or authorities). These findings corroborate our results, as those areas of science most familiar with RI tend to explore the subject further, and, consequently, are more likely to publish case details. Although the volume of published research in each research area also influences the number of cases, the predominance of medical and health sciences cases is not aligned with the trends regarding the volume of

published research. According to SCImago Journal & Country Rank (SCImago), natural sciences occupy the first place in the number of publications (41.5%), followed by the medical and health sciences (25.1%), engineering (22%), social sciences (7.8%), and the humanities (1.7%). Moreover, biomedical journals are overrepresented in the top scientific journals by IF ranking, and these journals usually have clear policies for research misconduct. High-impact journals are more likely to have higher visibility and scrutiny, and consequently, more likely to have been the subject of misconduct investigations. Additionally, the most well-known general medical journals, including NEJM, The Lancet, and the BMJ, employ journalists to write their news sections. Due to the fact that these journals have the resources to produce extensive news sections, it is, therefore, more likely that medical cases will be discussed.

Violations analysis

In the analysis of violations, the cases were categorized into major and minor misbehaviors. Most cases involved data fabrication and falsification, followed by cases involving non-adherence to laws and regulations, patient safety, plagiarism, and conflicts of interest. When classified by categories, 12.5% of cases involved issues in the study design, 16.4% in data collection, 56% in reporting, and 15.1% involved collaboration issues. Approximately 80% of cases involved serious research misbehaviors, based on the ranking of research misbehaviors proposed by Bouter and colleagues (2016). However, as demonstrated in a meta-analysis by Fanelli (2009), most self-declared cases involve questionable research practices. In the meta-analysis, 33.7% of scientists admitted questionable research practices, and 72% admitted, when asked about the behavior of colleagues. This finding contrasts with an admission rate of 1.97% and 14.12% for cases involving fabrication, falsification, and plagiarism. However, Fanelli's meta-analysis does not include data about research misbehaviors in its wider sense but focuses on behaviors that bias research results (i.e., fabrication and falsification, intentional non-publication of results, biased methodology, misleading reporting). In our study, detrimental research practices made up only 4.1% of the cases, while FFP made up 66.4%. Overrepresentation of some types of violations, and underrepresentation of others, might lead to misguided efforts, as cases that receive intense publicity eventually influence policies relating to scientific misconduct and

RI (National Academy of Sciences, Institute of Medicine Panel on Scientific, and the Conduct of Research 1992).

Sanctions analysis

The five most prevalent outcomes were paper retraction, followed by exclusion from funding applications, exclusion from service or position, dismissal and suspension, and paper correction. This result is similar to that found by Redman and Merz (2008), who collected data from misconduct cases provided by the ORI. Moreover, their results show that fabrication and falsification cases are 8.8 times more likely than others to receive funding exclusions. Such cases also received, on average, 0.6 more sanctions per case. Punishments for misconduct remain under discussion, ranging from the criminalization of more serious forms of misconduct to social punishments (Bülow and Helgesson 2019), such as those recently introduced by China (Cyranoski 2018). The most common sanction identified by our analysis - paper retraction – is consistent with the most prevalent types of violation, that is, falsification and fabrication.

Publicizing scientific misconduct

The lack of publicly available summaries of misconduct investigations makes it difficult to share experiences and evaluate the effectiveness of policies and training programs. Publicizing scientific misconduct can have serious consequences and creates a stigma around those involved in the case. For instance, publicized allegations can damage the reputation of the accused even when they are later exonerated (Bauchner et al. 2018). Thus, for published cases, it is the responsibility of the authors and editors to determine whether the name(s) of those involved should be disclosed. On the one hand, it is envisaged that disclosing the name(s) of those involved will encourage others in the community to foster good standards. On the other hand, it is suggested that someone who has made a mistake should have the right to a chance to defend his/her reputation. Regardless of whether a person's name is left out or disclosed, case reports have an important educational function and can help guide RE- and RI-related policies (Bird 2004). A recent paper published by Gunsalus (2019) proposes a three-part approach to strengthen transparency in misconduct investigations. The first part consists of a checklist (Gunsalus, Marcus, and Oransky 2018).

The second suggests that an external peer reviewer should be involved in investigative reporting. The third part calls for the publication of the peer reviewer's findings.

Limitations

One of the possible limitations of our study may be our search strategy. Although we have conducted pilot searches and sensitivity tests to reach the most feasible and precise search strategy, we cannot exclude the possibility of having missed important cases. Furthermore, the use of English keywords was another limitation of our search. Since most investigations are performed locally and published in local repositories, our search only allowed us to access cases from English-speaking countries or discussed in academic publications written in English. It is also important to note that the lack of information from the extracted case descriptions is a limitation that affects the interpretation of our results. In our review, only 25 retraction notices contained sufficient information that allowed us to include them in our analysis in conformance with the inclusion criteria. Although our search strategy was not focused specifically on retraction and misconduct notices, we believe that if sufficiently detailed information was available in such notices, the search strategy would have identified them.

CONCLUSION

Case descriptions found in academic journals are dominated by discussions regarding prominent cases and are mainly published in the news section of journals. Our results show that there is an overrepresentation of biomedical research cases over other scientific fields when compared with the volume of publications produced by each field. Moreover, published cases mostly involve fabrication, falsification, and patient safety issues. This finding could have a significant impact on the academic representation of ethical issues for RE and RI. The predominance of fabrication and falsification cases might diverge the attention of the academic community from relevant but less visible violations and ethical issues, and recently emerging forms of misbehaviors.

CHAPTER 3 – CHARACTERIZATION OF PHD STUDENTS AND DISCIPLINARY DIFFERENCES ON PERCEPTIONS AND ATTITUDE TOWARDS RCR[5]

BACKGROUND

A growing body of literature has shown the importance of taking steps to curb questionable research practices and cases of misconduct. More than just isolated cases, questionable research practices have been identified as one of the causes responsible for the lack of reproducible scientific results (Baker 2016).

Questionable research practices (QRPs), also known as detrimental research practices, can be defined as "*actions that violate traditional values of the research enterprise and that may be detrimental to the research process*" (National Academies of Sciences and Medicine 2017). These practices involve not only individual researchers but also institutions, sponsors, and journals. QRPs examples include selective reporting, bad authorship and citation practices, conflicts of interest, overinterpretation of study results, and p-hacking.

The recommendations to foster integrity involve the various stakeholders in the research enterprise. It includes creating and improving practices and policies to address research misbehavior and the emerging threats to scientific integrity, and promote responsible conduct of research (RCR) (Mejlgaard et al. 2020). However, since research performing organizations (RPO) have an essential role in promoting and educating on research integrity, they are a central player in this endeavour (Forsberg et al. 2018). The institutional responsibility is emphasized in The European Code of Conduct for Research Integrity: "*research institutions and organizations develop appropriate and adequate training in ethics and research integrity to ensure that all concerned are made aware of the relevant*

[5] This chapter is partially based on the INTEGRITY project deliverable D2.2- Final report on questionnaire and qualitative interviews, available at https://h2020integrity.eu/wp-content/uploads/2021/08/D2.2-Final-report-on-questionnaire-and-qualitative-interviews.pdf

codes and regulations" (ALLEA 2017). Moreover, they should identify their weaknesses and address them by maintaining high standards for responsible conduct of research.

Over the past years, several initiatives from RPOs emerged to foster academic and scientific integrity using different methodologies and approaches (Mejlgaard et al. 2020). Regardless of the chosen approach, accumulated evidence shows that it is imperative to understand the context-dependent differences and tailor the interventions to ensure effectiveness (Viđak et al. 2021, Ščepanović et al. 2021).

Therefore, this study aims to characterize the PhD students in Hungary regarding perceptions and practices on RCR, more focused on the citation, collaboration, authorship, and data practices, and assess and understand the disciplinary differences. In addition, the study also aims to assess the prevalence of training in RCR and current learning methods.

METHODS

This study is part of the INTEGRITY project (https://h2020integrity.eu/)[6], a European Union's Horizon 2020 funded project that aims to map academic integrity in Europe across high school and university students. The questionnaire-based study recruited participants from seven countries, but we are focusing only on the results from PhD students in Hungary for this study.

The study was conducted in accordance with the ethical guidelines of research in Hungary.[7] The participants provided informed consent at the beginning of the survey.

Sampling and data collection

All Doctoral Schools in Hungary were randomized and stratified according to three main fields: 1- Science, Technology, Engineering, and Mathematics (STEM); 2- Social Sciences (SOCSCI); 3- Humanities (HUM). Firstly, based on a list of universities in Hungary, the

[6] INTEGRITY project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement N824586. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

[7] As the expected negative impact does not exceed the daily emotional experiences, ethics approval is not required under regulations in Hungary.

universities with programs in the three areas were randomized. Then, all doctoral schools of the selected universities were randomized, selecting five doctoral schools from each area. The doctoral schools were then contacted for participation. The PhD offices sent the survey link to all PhD students who completed at least the first year. The recruitment process was repeated until reaching at least 50 participants from each area. Ninety-two doctoral schools were contacted (40 STEM, 29 SOC SCI, 23 HUM). Fifty agreed to participate (22 STEM, 17 SOCSCI, 11 HUM). The data were collected between February and December 2020.

Instrument

The survey was developed as part of the INTEGRITY project. The content and the addressed topics were based on a qualitative interview study, also part of the same project (Goddiksen et al. 2021). The instrument was validated in a pilot study. The pilot questionnaire was applied to the relevant student population in The Netherlands, Portugal, and Lithuania using an online survey platform. The tests and preliminary analysis were conducted and changes to the survey tool were implemented based on pilot data. The questionnaire was developed in English and translated to Hungarian. The survey was an anonymous online survey using SurveyXact ver. 12.9 (https://www.surveyxact.com/).

The questionnaire included demographic questions (gender, age, level of study, and country). The study also included questions on whether the participants collaborate or plan to collaborate with others as part of their research and what type of data they collect or analyze in their research. These questions help tailor the following questions according to their research context. For instance, a participant who selected quantitative data would get a statement "Not mentioning in a research publication that you replaced some outliers in a data set with data points obtained through estimates based on the remaining data points". While other who selected interviews as their type of data would get "Not mentioning in a research publication that you made up a number of direct quotes from an interview based on your general impression of the interviewee's viewpoints". See Appendix 4 for the full description.

The study's specific questions addressed research integrity training and learning methods, good research practices understandings, self-reported attitude, and observed peer's attitude.

The questions included situations in the three explored themes (plagiarism and citation practices, collaborations and authorship practices, and data practices.

Under the training and education questions, the participants were asked whether they had any training on academic integrity, and the answers included “Yes, one or more dedicated courses”, “Yes, one or more lectures”, “Yes, one or more dedicated e-learning sessions”, or “No”. The respondents were asked if they have learned about RCR through any other method, such as supervisors and teachers, fellow students, self-study, or others. They were also asked about the most important topics on RCR in their field of study.

Under the “understanding of good research practices” questions, they were asked whether they believe they have a good understanding of the official standards of good practice in relation to the three explored themes mentioned above. The answers could range from “Fully agree” to “Fully disagree”, and “I don’t know”. They were also asked if, over the past 12 months, they were in a situation where they were unsure how to behave in an ethically correct manner in relation to the three themes. The potential answers were: “Yes, many times”, “Yes, a few times”, “Yes, once”, “No”, and “Not applicable”. Then, they were introduced to some situations on the three explored themes and asked whether they believed the following actions go against the rules and regulations. The scenarios included clear misconduct, questionable research practices and grey zones. The answers could range from “Yes, it is a serious violation” to “No, it is not against the rules”, including “The rules are unclear”, ‘It depends on the situation”, and “I don’t know”.

When assessing attitudes, the PhD students were introduced to some scenarios with misconduct and QRPs, and they were asked to what extent they agreed with the statement “It is common for researchers in my field to…”. The answers ranged from “Fully agree” to “Fully disagree”, and “I don’t know”. They were later introduced to the same scenarios and asked whether they have done something similar. The potential answers were: “Yes, many times”, “Yes, a few times”, “Yes, once”, “No”, “I prefer not to answer”, “Not applicable”, and “I don’t know”. They were also asked why they think their peers deviate from good practices. Those who reported having added as coauthor “someone who did not make a significant contribution to the research publication” on the question of self-reported attitude were also asked about their motivations to do so.

Data analysis

The analyses were performed using IBM SPSS version 28.0 and Stata 17.0. The descriptive results for all cases were reported. In order to analyze the disciplinary differences, the Kruskal-Wallis was used for ordinal data with more than two groups, and Mann-Whitney U test for two groups of comparison. Pearson's Chi Squared and Fisher's exact tests were used for nominal data. The effect sizes for the ordinal data were calculated using eta square, where $\eta^2=H\text{-}k+1/n\text{-}k$, and the interpretation values are: 0.01- < 0.06 (small effect), 0.06 - < 0.14 (moderate effect) and >= 0.14 (large effect). The effect sizes for the nominal data were calculated using Cramer's V. The interpretation for Cramer's V is based on the degrees of freedom. The degrees of freedom are defined by df = min(#rows-1, #columns-1), and in this study, the degrees of freedom were 3 in all the analyses. Therefore, the interpretation values commonly in published literature are: 0.06- < 0.16 (small effect), 0.17 - < 0.28 (moderate effect) and >= 0.29 (large effect) (Cohen 2013).

The "I don't know" answers were excluded from the analyses of Self-perception on knowledge and understanding of good practices, and Perceptions on peers' attitude towards RCR. Where the options "The rules are unclear" and "It depends on the situation" were available, the answers were coded the same. The aswers "Not applicable" were excluded of the Self-reported ethical uncertainty analyses. Multivariate regression analyses were performed on significant variables in univariate analysis to assess confounding interactions. The statistically significant associations were considered at the 0.05 level.

RESULTS

Approximately 2082 students received the invitation to participate from the selected doctoral schools. A total of 200 PhD students completed the survey, giving an approximately 10% response rate. Eighty-three (41.5%) respondents were male and 117 (58.5%) were female. Most of the PhD students are from STEMM field (50%), followed by Social Sciences (25%), Humanities (21.5%), and other disciplines (3.5%).

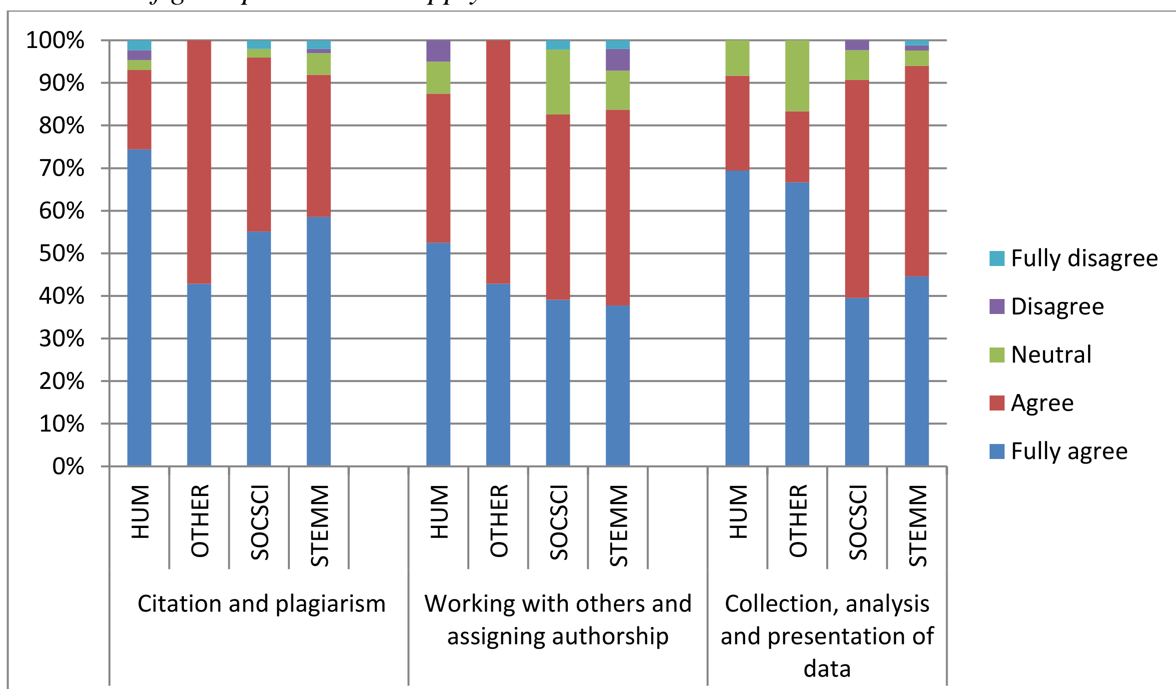
1. Self-perception on knowledge and understanding of good practices

The questionnaire included three questions on self-perception of good practices understandings. One on citation and plagiarism, one regarding working with others and authorship practices and collection, and one on analysis and presentation of data. The results show that citation and plagiarism had the highest prevalence of positive answers (Fully agree and agree) and authorship practices had the lowest (Table 1). Although there were some divergences across scientific fields, a Kruskal-Wallis test showed no significant differences ($p=0.270$, $\eta^2=0.00$; $p=0.454$, $\eta^2=0.00$; and $p=0.078$, $\eta^2=0.02$ for citation and plagiarism, authorship practices, and data practices, respectively) (Figure 1).

Table 1. Self-perception on knowledge and understanding of good practices

	Citation and plagiarism	Authorship practices	Data practices
Fully agree	61%	41%	49%
Agree	33%	43%	43%
Neutral	4%	10%	6%
Disagree	1%	4%	1%
Fully disagree	2%	2%	1%

Figure 1. Responses to the statement: "*I have a good understanding of the official standards of good practice that apply to me in relation to…*"



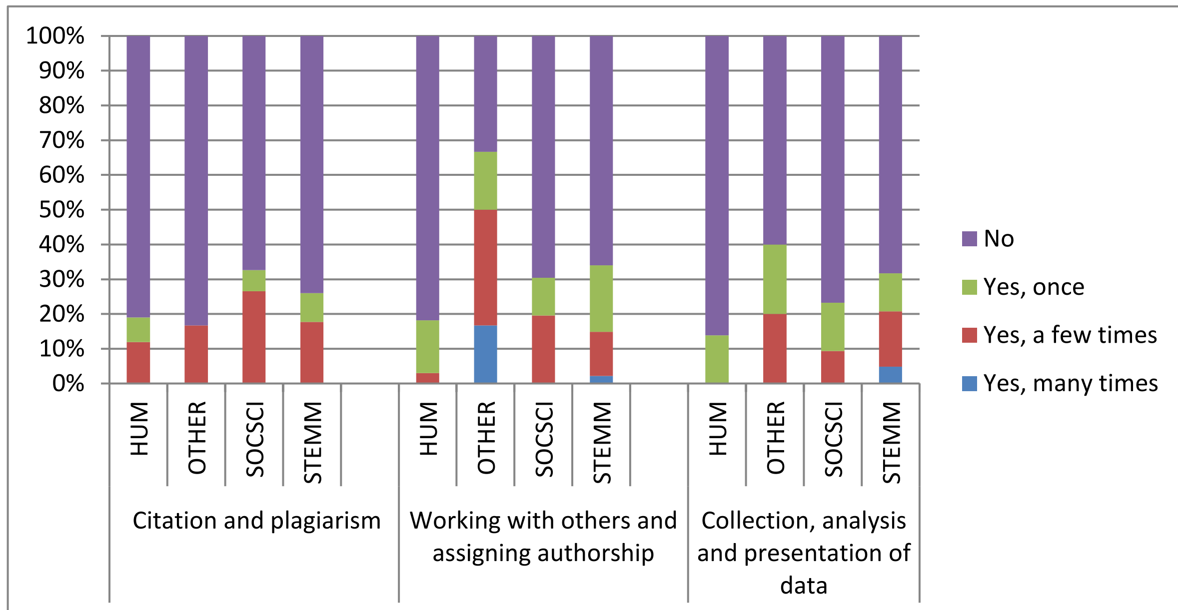
2. Self-reported ethical uncertainty

The students were asked whether they had experienced uncertainties about how to behave in relation to the three explored themes in the past 12 months. The results show that, in general, collaboration and authorship practices were the most frequent topic causing uncertainties (31%) (Table 2). A Kruskal-Wallis test showed that disciplinary differences significantly affect uncertainty when facing situations involving collaboration and authorship practices ($p=0.038$, $\eta^2=0.03$) (Figure 2), even after adjusted for confounding (gender and RI course attendance). For the other two themes, approximately 26% of the respondents reported having experienced doubts about how to behave in situations involving citation and plagiarism, and data practices. However, no significant differences were found across disciplines ($p=0.430$, $\eta^2=0.00$, and $p=0.096$, $\eta^2=0.02$ for citation and plagiarism and data practices, respectively.

Table 2. Self-reported ethical uncertainty

	Citation and plagiarism	Authorship practices	Data practices
Yes, many times	0%	2%	2%
Yes, a few times	19%	13%	11%
Yes, once	7%	16%	13%
No	74%	69%	74%

Figure 2. Responses to the statement: "*Over the past 12 months, have you been in a situation where you were unsure how to behave in an ethically correct manner in relation to …*"



3. Perceptions and practical judgments on rules of research practices

The questionnaire included questions on rules and good practices about the three explored themes. The questions included scenarios that are clearly regarded as misconduct and more complex QRPs. The answers included a range from "Yes, it is a serious violation" to "I don't know".

3.1 Plagiarism and citation practices

In the theme of plagiarism and citation practices, we introduced four scenarios:

Scenario 1 - Copying an entire page stating a central point from an external source into your own text without quotation marks but including a reference.

Scenario 2 - Copying one short paragraph a central point from an external source into your own text without quotation marks but including a reference.

Scenario 3 - Changing 10% of the words in a short paragraph stating a central point from an external source and using it in your own text with a reference.

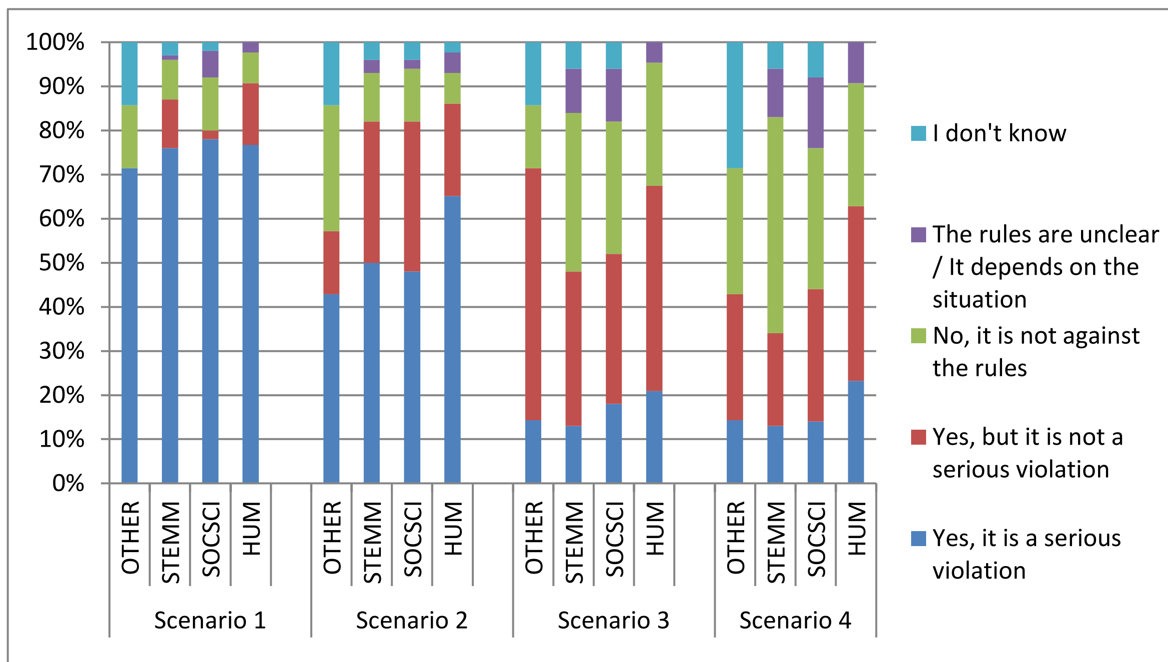
Scenario 4 - Copying a central point formulated in half a sentence from an external source without marking it with quotation marks but including a reference.

The results show that in the clear misconduct scenario (Scenario 1), the large majority (77%) of the respondents stated that copying an entire page without quotation marks is a serious violation (Table 3). In general, in Scenario 2, a majority also reported it as a "serious violation" (53%). However, despite being an unacceptable practice, there was an increase in classifications as a "non serious violation" (30%) and as an acceptable practice (11%). Scenario 3 displays a grey zone. In this case, although a majority classified it being a violation, a larger proportion reported as an acceptable practice (32%). When comparing the perceptions across fields, a Pearson's Chi-Squared test showed no significant differences (p=0.272, V=0.16; p=0.662, V=0.13; and p=0.566, V=0.13 for scenarios 1, 2, and 3, respectively). In Scenario 4, it is possible to see a difference between fields, as in Humanities over 60% of the respondents reported it as a violation, while 35% of STEMM respondents (Figure 3). A Pearson's Chi-Squared test showed significant disciplinary differences for the scenario 4 (p=0.044, V=0.19). However, when adjusted for confounding (integrity course attendance and gender), the difference did not remain significant.

Table 3. Perceptions and practical judgment on plagiarism and citation practices

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Yes, it is a serious violation	77%	53%	16%	16%
Yes, but it is not a serious violation	9%	30%	38%	28%
No, it is not against the rules	10%	11%	32%	40%
The rules are unclear / It depends on the situation	3%	3%	9%	12%
I don't know	3%	4%	5%	6%

Figure 3. Answers to the statement: “Please indicate whether you believe the following actions go against the official rules and regulations that apply to you in relation to plagiarism and citation practices”.



3.2 Collaboration and authorship practices

In the theme of collaboration and authorship practices, we introduced three scenarios:

Scenario 1 - Adding someone who has not written anything as an author of a research publication.

Scenario 2 - Adding a supervisor who only provided critical comments on the manuscript as a coauthor of the research publication.

Scenario 3 - Removing a coauthor who made a genuine contribution to the writing of a manuscript because the person does not want to be a coauthor.

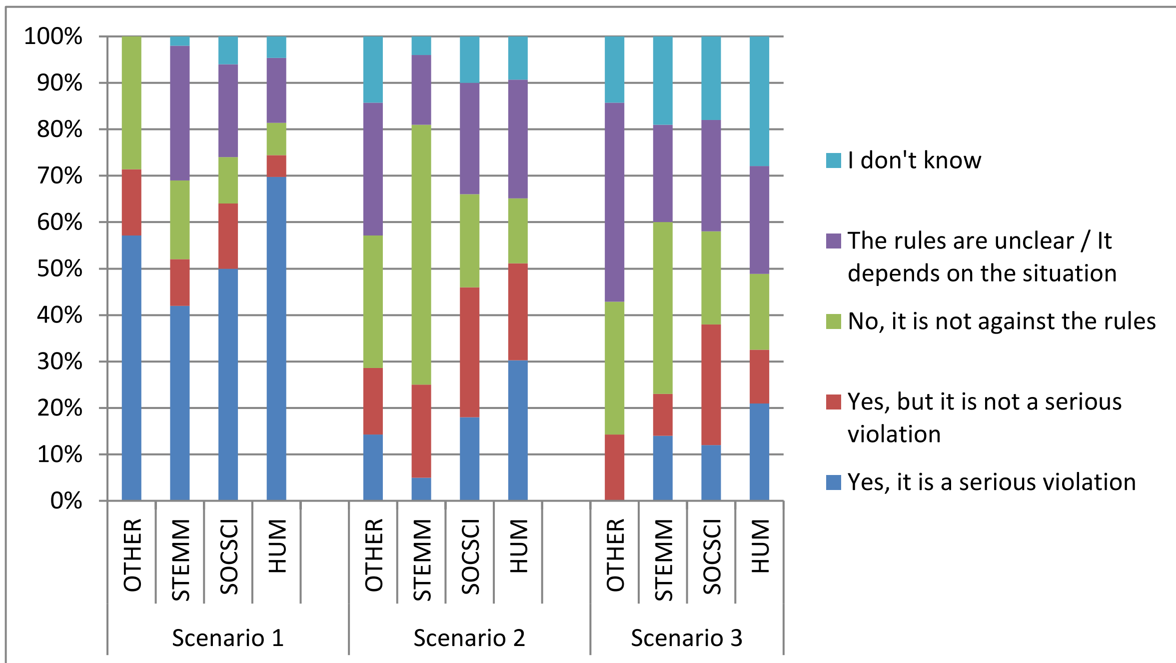
The results show that for scenario 1, only approximately 50% of STEMM PhD students perceive it as a violation to add someone as coauthor who does not qualify to be one. In contrast, approximately 75% of Humanities students perceive it in the same way (Figure 4).

However, a Pearson's Chi-Squared test showed no significant differences between fields (p=0.127, V=0.17). In scenario 2, the perceptions on the degrees of violation were statistically significant across disciplines (p<0.001, V=0.26), even after adjusted for confounding. Twenty-five percent of STEMM students perceive "Adding a supervisor who only provided critical comments on the manuscript as a coauthor of the research publication" as a violation, while 51% of Humanities students reported it as a violation. Students from Social sciences and other fields ranged between the other two groups. In scenario 3, it is possible to see differences between STEMM and Social sciences students, where 23% of STEMM students see the scenario as a violation, while 38% of social sciences students would classify the situation as a violation. However, no significant differences were found (p=0.097, V=0.18). In general, regardless of the scenario, a substantial proportion of PhD students could not deliberate about the cases and chose the options "I don't know" or "The rules are unclear/It depends on the situation" (Table 4). The proportion of uncertainties in these cases was larger for grey zones situations, 27%, 27%, and 44% for scenarios 1, 2, and 3, respectively.

Table 4. Perceptions and practical judgment on authorship practices

	Scenario 1	Scenario 2	Scenario 3
Yes, it is a serious violation	51%	14%	15%
Yes, but it is not a serious violation	10%	22%	14%
No, it is not against the rules	14%	37%	28%
The rules are unclear / It depends on the situation	23%	20%	23%
I don't know	4%	7%	21%

Figure 4. Answers to the statement: “Please indicate whether you believe the following actions go against the official rules and regulations that apply to you in relation to collaboration and authorship practices”.



3.3 Data practices

In the theme of data practices, we introduced four scenarios. The scenarios’ phrasing was tailored for each field and type of data. See Appendix 4 for the full description of all the cases. For those who worked with quantitative data, the scenarios were:

Scenario 1 - Not mentioning in a research publication that you removed some deviating data points from a data set when the cause of the deviation was unknown.

Scenario 2 - Not mentioning in a research publication that you removed some deviating data points from a data set when the cause of the deviation was known.

Scenario 3 - Not mentioning in a research publication that you replaced some outliers in a data set with data points obtained through estimates based on the remaining data points.

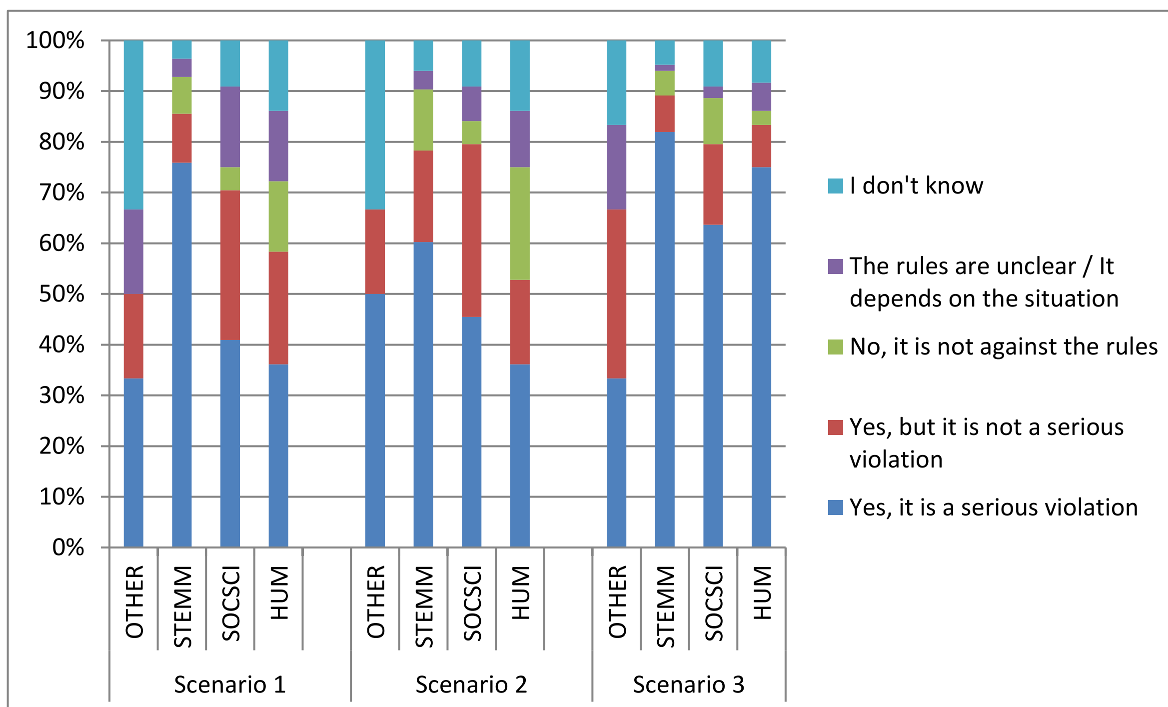
In scenario 1, the perceptions on the degrees of violation were statistically significant across disciplines according to Pearson’s Chi-Squared test ($p=0.001$, $V=0.26$), even after

adjusted for confounding (Figure 5). Eighty-five percent of STEMM students see the practice as a violation, while only 50% from other disciplines and 58% of Humanities students see it as a violation (Table 5). In scenario 2, when the cause of deviation was known, the differences were significant (p=0.044, V=0.21). The major differences are also between STEMM and Humanities, where 78% and 52%, respectively, see the practice as a violation. However, the difference did not remain significant after adjusting for confounders. In scenario 3, although there were small differences when perceiving the severity of the violation, no significant differences were found (p=0.147, V=0.18).

Table 5. Perceptions and practical judgment on data practices

	Scenario 1	Scenario 2	Scenario 3
Yes, it is a serious violation	57%	51%	74%
Yes, but it is not a serious violation	18%	22%	11%
No, it is not against the rules	8%	12%	5%
The rules are unclear / It depends on the situation	9%	6%	3%
I don't know	8%	9%	7%

Figure 5. Answers to the statement: "Please indicate whether you believe the following actions go against the official rules and regulations that apply to you in relation to data practices".



4. Perceptions on peers' and self-reported attitude towards RCR

The questionnaire included questions on the respondents' perceptions of peers' and self-reported attitudes towards RCR. The questions included examples of clear misconduct and QRPs in relation to the three explored themes.

4.1 Perceptions on peers' attitude towards RCR

The respondents were asked to what extent they agreed with the following statement: "It is common for researchers in my field to...". The answers range from "Fully agree" to "Fully disagree". The scenarios' phrasing was tailored for each field and type of data. See Appendix 4 for the full description of all the cases. For those who worked with quantitative data, the scenarios were:

Scenario 1 - Allow research group leaders, supervisors or others in power to become coauthors of papers, even though they did not make a significant contribution to them.

Scenario 2 - Delete deviating data points based on a gut feeling that they were inaccurate.

Scenario 3 - Perform misleading or dubious statistical analysis in order to achieve publishable results.

Scenario 4 - Keep inadequate records of parts of their research that should be documented (e.g., laboratory records, descriptions of data sampling procedures, key words used in a literature study, procedures for qualitative coding, types of statistical tests performed etc.).

Scenario 5 - Copy shorter passages from other sources into their research papers without marking them as quotes.

Scenario 6 - Cite sources that are not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour.

In general, among the presented scenarios, scenario 1 had the highest positive agreements (Fully agree and agree) with 46% of the respondents, followed by scenario 6 (32%), scenario 4 (24%), scenario 2 (23%), scenario 3 (17%), and scenario 5 (13%) (Table 6).

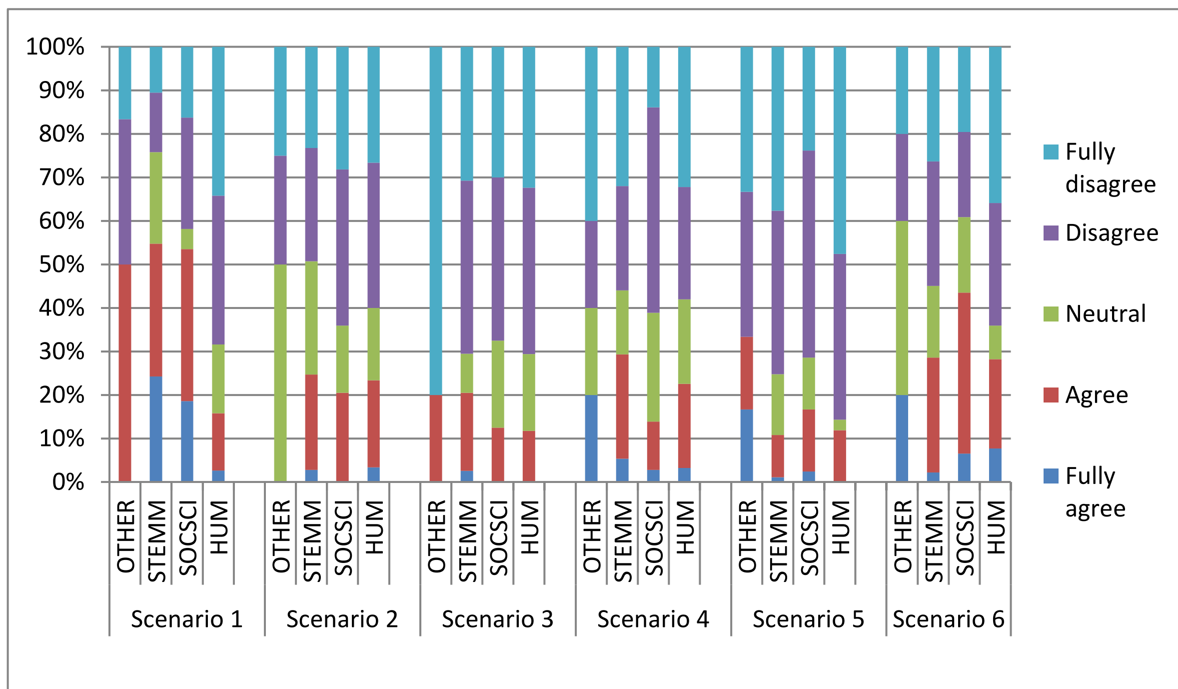
Table 6. Prevalence of the perceptions on peers' attitude for each presented scenario

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Fully agree	18%	2%	1%	5%	2%	5%
Agree	29%	21%	15%	19%	12%	27%
Neutral	15%	22%	13%	18%	10%	16%
Disagree	21%	30%	38%	30%	40%	26%
Fully disagree	17%	25%	33%	28%	37%	27%

In scenario 1 (Figure 6), a Kruskal-Wallis test showed a statistically significant difference in the respondents' perceptions of their peers' attitudes regarding authorship ($p<0.001$, $\eta^2=0.14$). Post-hoc test Bonferroni-corrected shows significant differences between the Humanities and Social Sciences ($p=0.007$) and between Humanities and STEMM ($p<0.001$). The associations remained significant after adjusting for confounding. Humanities PhD students are less likely to observe their colleagues assigning authorship to someone who does not qualify to be one. In scenario 2, 21-25% of the three major study areas agreed that their peers arbitrarily delete deviating points. There were no significant

differences across areas ($p=0.687$, $\eta^2=0.01$). In scenario 3, the three major study areas have similar perceptions (approximately 30%) on their peers' attitudes regarding misleading interpretations or analysis ($p=0.469$, $\eta^2=0.00$). In scenario 4, although Social sciences students were less likely to agree to the statement about inadequate data management, there were no differences across fields ($p=0.969$, $\eta^2=0.02$). In scenario 5, there were no differences across fields regarding the respondent's perceptions of the peers' attitudes about plagiarism and citation practices ($p=0.140$, $\eta^2=0.01$). In scenario 6, although Social sciences students were more likely to agree to the statement about bad citation practices, there were no differences across fields ($p=0.181$, $\eta^2=0.01$).

Figure 6. Answers to the statement: "It is common for researchers in my field to...".



4.2 Self-reported attitude towards RCR

The questionnaire included questions about the self-reported attitudes towards RCR. The respondents were introduced to the same six scenarios. In this case, they were questioned whether they had done any of the scenarios (During your PhD, have you...). The answers included "Yes, many times", "Yes, a few times", "Yes, once", "No", or "I prefer not to answer". See Appendix 4 for the full description of all the cases. For those who worked with quantitative data, the scenarios were:

Scenario 1 - Allowed research group leaders, supervisors or others in power to become coauthors of papers, even though they did not make a significant contribution to them.

Scenario 2 - Deleted deviating data points based on a gut feeling that they were inaccurate.

Scenario 3 - Performed misleading or dubious statistical analysis in order to achieve publishable results.

Scenario 4 - Kept inadequate records of parts of their research that should be documented (e.g., laboratory records, descriptions of data sampling procedures, key words used in a literature study, procedures for qualitative coding, types of statistical tests performed etc.).

Scenario 5 - Copied shorter passages from other sources into their research papers without marking them as quotes.

Scenario 6 - Cited sources that are not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour.

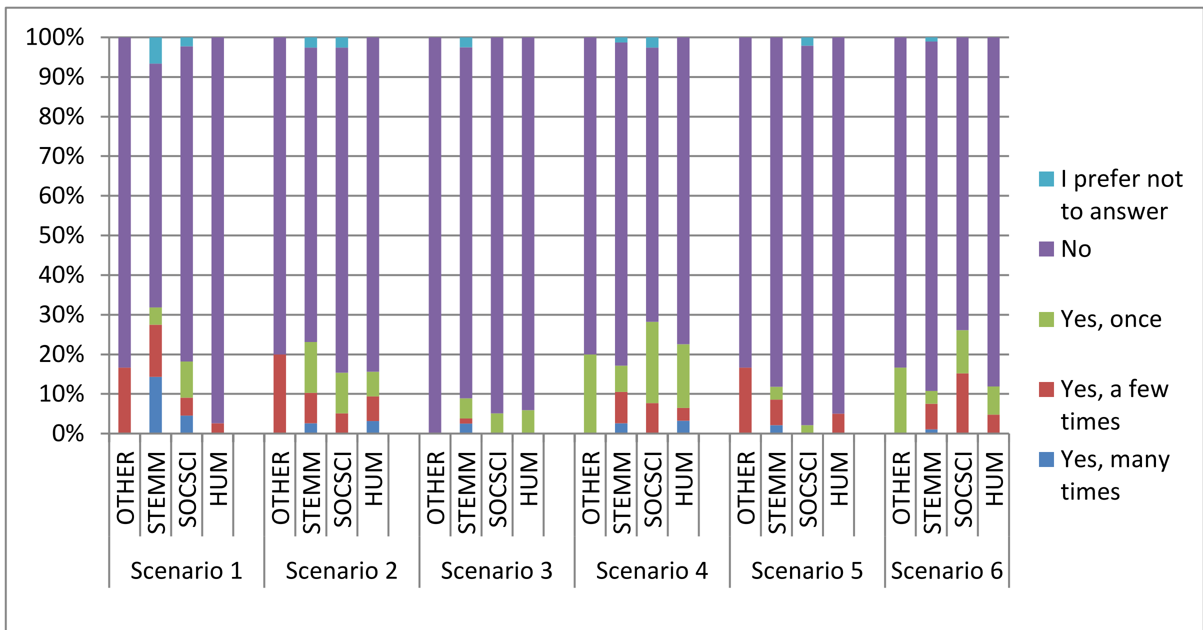
In general, among the presented scenarios, scenario 1 was also the most self-reported practice with 22% of the respondents, followed by scenario 4 (21%), scenario 2 (19%), scenario 6 (15%), scenario 5 (8%), and scenario 3 (7%) (Table 7).

Table 7. Prevalence of self-reported attitude for each presented scenario

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Yes, many times	8%	2%	1%	2%	1%	1%
Yes, a few times	9%	7%	1%	7%	5%	8%
Yes, once	4%	10%	5%	13%	2%	6%
No	74%	79%	92%	77%	91%	84%
I prefer not to answer	4%	2%	1%	1%	1%	1%

In scenario 1, as in perceptions of peers' attitudes, the self-reported attitude had significant differences across disciplines, according to Fisher's exact test (p=0.007, V=0.22), even after adjusted for confounding. While 32% of STEMM PhD students reported having assigned authorship to someone who does not qualify to be one, only 3% of Humanities PhD students reported the same (Figure 7). For all the other scenarios, there were no significant disciplinary differences (p=0.92, V=0.11; p=0.992, V=0.11; p=0.529, V=0.14; p=0.337, V=0.14; and p=0.281, V=0.14, for scenarios 2, 3, 4, 5, and 6, respectively).

Figure 7. Answers to the statement: “During your PhD, have you...”.



The PhD students who reported having added as coauthor someone who does not qualify to be one were asked about their main motivations to do so. The most prevalent answers were because “I wanted to maintain a good relationship with the person” (34%), “The person in power told me to” (28%), and “Everyone else in my field does it” (20%) (Table 8). A Kruskal-Wallis test with a Bonferroni-corrected post-hoc showed significant differences between STEMM and Humanities when stating the motivation “The person in power told me to” ($p=0.018$, $\eta^2=0.05$). There were also significant differences when reporting as motivation “Everyone else in my field does it” between STEMM and Social sciences ($p=0.029$, $\eta^2=0.05$). The associations remained significant after adjusting for confounding.

Table 8. Reasons to add a coauthor that does not qualify to be one

	OTHER	STEMM	SOCSCI	HUM
The person in power told me to	0%	20%	6%	2%
I feared I would not be awarded my degree if I didn’t	0%	3%	0%	0%
Everyone else in my field does it	0%	16%	2%	2%
Friends and/or family encouraged it directly or indirectly	0%	0%	0%	0%
I believed they deserved it	0%	6%	2%	0%
I wanted to maintain a good relationship with the person	14%	13%	4%	2%
Other reasons	0%	3%	2%	0%
I prefer not to answer	0%	1%	0%	0%

5. Education on RCR

The questionnaire included questions regarding training and education on RCR. The PhD students were asked whether they had taken any course on rules and/or ethically correct behavior in relation to the three main themes during their current or previous studies. In general, 62.5% reported having attended a course on the related topics. They were also asked how these courses were delivered. The options were dedicated courses, lectures, or e-learning sessions, and they could select more than one option. The most selected option was "One or more lectures" (45.5%), followed by "One or more dedicated courses" (26.5%), and "One or more dedicated e-learning sessions" (5%). There were no significant disciplinary differences in course attendance or type of course.

The PhD students were also asked if they had learned about the themes and topics through any other method (Table 9). The most selected option was through "Supervisors/teachers in other courses that commented on my written work or assignments", followed by "Following the procedures that are common in my field of study", and "Discussions with teachers outside regular classes".

Table 9. Used methods to learn abour RCR topics

	Total	HUM	OTHER	SOCSCI	STEMM	p value
Supervisors/teachers in other courses that commented on my written work or assignments	47%	70%	14%	46%	40%	**0.003**
Courses not dedicated exclusively to such issues	32%	44%	14%	38%	25%	0.071
Discussions with fellow students	34%	40%	57%	30%	32%	0.425
Discussions with teachers outside regular classes	40%	49%	14%	42%	37%	0.286
Self-study	38%	37%	43%	44%	34%	0.680
By following the procedures that are common in my field of study	42%	58%	29%	34%	39%	0.078
Discussions with friends and family outside my institution	10%	16%	0%	4%	10%	0.185
Other	9%	9%	29%	8%	7%	0.265

DISCUSSION

This study aimed to characterize the PhD students in Hungarian universities regarding perceptions and attitudes of RCR and identify potential differences between scientific fields. The study mainly focused on three themes, plagiarism and citation practices, collaboration and authorship practices, and data practices. According to our results, PhD students' self-evaluation regarding their understanding of RCR was good. However, when confronted by practical situations, a fraction of the respondents was unable to identify violations in misconduct cases and QRPs.

Understanding and perceptions of QRPs

When assessing plagiarism and citation practices, over 90% of the PhD students reported having a good understanding of rules and regulations. However, a quarter of the students was unable to classify a clear case of plagiarism (copying an entire page but including a reference in the text) as a serious violation. The proportion increased when the scenario was modified by the length of the used text (a whole page was changed for a short paragraph), where 47% did not see it as a serious violation.

The uncertainties concerning the concept of plagiarism have been largely documented. Students most often know when to reference but rarely know exactly how. The lack of detailed knowledge on regulations and conventions about paraphrasing and quoting often lead to unintentional plagiarism (Park 2003). The common misunderstanding for authors is that changing a few words to avoid quotation is acceptable, although essentially identical. Moreover, the concept's disagreements tend to be even greater regarding seriousness and degrees of violation (Gullifer and Tyson 2014). Although disciplinary differences can affect the perceptions of what constitutes plagiarism and improper paraphrasing, no significant differences were found in this study.

Collaboration and authorship practices had the lowest rate of understanding among the three explored themes. Eighty-four percent of the PhD students reported a good understanding of the standards. However, despite this self-reported understanding, only 49% of the respondents adequately classified "assigning authorship to someone who has

not written anything" as a serious case of violation. When asked about "assigning authorship to a supervisor who only provided critical comments", only 36% perceived it as a violation. Authoring scientific papers is fundamental for career advancement and prestige in modern science. On the other hand, authorship policies of different scientific journals and publishers can be dubious and are frequently revised. There are no universal standards on authorship, and the regulations vary between scientific disciplines. For instance, in the Natural sciences, the guidelines of the PNAS (Proceedings of the National Academy of Sciences of the United States of America) are often indicated. In contrast, the APA guidelines (American Psychological Association) are more frequently used in the Social Sciences. The guidelines of ICMJE (International Committee of Medical Journal Editors) are well-known and probably the most prestigious. This document lists four authorship criteria, which must be all fulfilled to qualify a contributor as an author. Those who do not meet all four criteria should be mentioned in the acknowledgement section. Regardless of the field, the authors need to understand and adhere to the criteria used in their own discipline. The common minimum requirements to all definitions are a substantial contribution to the work and accountability for the work.

When comparing the fields, there was a significant difference between Humanities and STEMM regarding the scenario involving adding a supervisor who only provided critical comments on the manuscript as a coauthor. In this case, most STEMM students perceive it as a non-violation, while most students from the Humanities perceive it as a serious violation. Despite the disciplinary differences, this scenario would constitute a violation of authorship guidelines for both fields. In the Humanities field, publications are usually single-authored and with fewer coauthors than in other areas (Henriksen 2016). Thus, unfair assignment of authorship might be less frequent, and the rules might be more transparent or less relevant. These findings were corroborated in a recent study investigating disciplinary differences in QRPs where Humanities respondents were the only ones not discussing practices of unfair authorship assignment (Ravn and Sørensen 2021).

Over 90% of the PhD students reported having a good understanding of data practices. However, the results largely varied across disciplines when asked about the degrees of violation or seriousness of some data practices. Humanities students are less likely to

perceive "deletion of data, interview, or material that did not fit well with the rest of your material when the cause of the deviation was unknown" as a violation than those of STEMM fields. The same applies to situations where the deviation is known.

In a study by de Vries, Anderson, and Martinson (2006), the results show that researchers are often in doubt about the rules and standards applied of RCR. This study found that data handling and cleaning were the most visible difficulties when interpreting the rules. Researchers reported uncertainties about the line between data cleaning and data manipulation. Although intentional data manipulation can occur (Fanelli 2009), understanding statistical reasoning, data analysis, and interpretation are difficult, and error, lack of knowledge, and a bad learning environment can lead to unintentional QRPs (Sijtsma 2016).

Peer and self-reported behavior

The students were asked about peers' and self-behavior. In general, regarding plagiarism and citation practices, 32% observed and 15% admitted having cited an article to please someone. Although this topic is not explored in most surveys, it was ranked 5 in the most frequent misbehavior (Bouter et al. 2016). Furthermore, 14% of the PhD students reported that improper quotation is a common practice for researchers in their field, and 6% admitted having done the same at least once. An improper quotation can be characterized as a form of plagiarism (Gullifer and Tyson 2014). Therefore, if we compare to survey results on plagiarism, our results are relatively lower than those found in Croatia (29.1% and 9.3%), and to the observed results in a meta-analysis (30% CI 17%-46%) (Pupovac and Fanelli 2015b).

Regarding collaboration and authorship practices, in general, 47% of the PhD students reported that "allowing research group leaders, supervisors or others in power to become coauthors of papers, even though they did not make a significant contribution to them is common for researchers in their field. These results are similar to those found in Croatia (46%), and higher than in Scandinavian countries (19%). However, only 21% admitted having done the same. A meta-analysis on authorship (Marušić, Bošnjak, and Jerončić 2011) shows that 29% (95% CI 24%-35%) observed or self-reported misusing authorship.

Although our results show misunderstandings on authorship assignment, fair authorship assignment is not only a matter of knowing the existence of guidelines, but more importantly, is putting them into practice. This requires a culture where mentors and fellow colleagues act as role models for early-career researchers. However, in our study, most of the respondents who admitted having added someone who does not qualify as coauthor responded that they have done so because the person in power told them to. They wanted to maintain a good relationship with the person, and because everyone in their field does it. These motivations are not surprising, as PhD students feel under pressure to add authors when required by their supervisors. In a Danish study, 32% of the PhD students believed they were under pressure regarding the inclusion or order of authors (Jensen et al. 2018). The researchers often report having to "negotiate standards of conduct", balancing the meaning of data, rules, relation with colleagues, and productivity pressure (de Vries, Anderson, and Martinson 2006). Science is a social activity, and in order to maintain a good relationship with others, researchers have to make trade-offs between scientific authorship or other benefits.

When asked about peers' behavior on data practices, 23% reported that it is common practice for researchers in their field to delete deviation points, or material, perform misleading data analysis or interpretation (16%), and inadequate records keeping about their research (24%). These results contrast to those in Croatia (Ljubenković et al. 2021), where misleading results was the most reported misbehavior. Nineteen percent admitted the deletion of deviation points or material, 7% had performed misleading data analysis or interpretation, and 22% had kept inadequate records of their research.

For the three explored themes of our study, disciplinary differences were significant only regarding authorship practices. When comparing disciplines, STEMM students were more likely to report adding someone who does not qualify as coauthor as common in their field than those from Humanities. The same applied to self-reported behavior. This difference can be explained by the different disciplinary and publication cultures in the Humanities field, as already mentioned in the perceptions of authorship practices.

In general, there was a substantial difference between the self-admission and the observed proportions for all the investigated violations. As demonstrated by prevalence studies on

misconduct and QRP's (Pupovac and Fanelli 2015a, Xie, Wang, and Kong 2021), respondents are less likely to self-report questionable behaviors than to report these behaviors as being observed on others. The self-reported behaviors are usually underestimated due to the stigmatization of QRPs. However, it is hard to predict if the observed behaviors are under or overestimated. Therefore, these findings shed light on the prevalence of misconduct and QRPs in Hungary, as PhD students are more likely to underestimate research misbehaviors.

Education

Sixty-two percent of the PhD students had attended a course on RCR. They were delivered mainly by lectures, followed by courses, and only a few e-learning sessions. There were no questions regarding the compulsory status of the training. These findings corroborate those found in European universities, where RCR education is not available in all institutions in Europe (Abdi et al. 2021). This recent study investigating research integrity education in Europe had few recommendations. Firstly, as found in all the current recommendations, institutions have the duty to promote RCR instructions (Mejlgaard et al. 2020, ALLEA 2017). They also recommend that European universities set minimal requirements for research integrity training, as specified in the US. In the US, the NIH recommends "at least 8 contact hours of face-to-face discussions, recurring at least once during each career stage, and at a frequency of no less than once every four years".

Limitations

The main limitation of our study is the response rate. Only 10% of invited PhD students completed the survey. Due to data protection, we could not directly contact the students, and the invitations were sent via Doctoral Schools. Therefore, sending reminding e-mails was not an option. To minimize the impact of the response bias, we recruited participants until reaching a minimum of fifty respondents from each broad scientific field. Moreover, the survey phrasing could have affected the understanding for some fields in Hungary, despite having tailored the questions according to the chosen scientific background.

CONCLUSION

Overall, our results show a discrepancy between the self-reported understanding and the perception and knowledge of rules on RCR. The PhD students in Hungary believe that they have a better understanding of RCR standards than in practice. Unfair authorship assignment and data handling are the most causes of uncertainties. Overall, disciplinary differences were significant only for authorship practices. Our results also show a high prevalence of admitted and observed QRPs. Thus, it highlights the importance of promoting research integrity through effective training and fostering an ethical research culture.

CHAPTER 4 – PERCEPTIONS OF RESEARCH INTEGRITY CLIMATE IN HUNGARIAN UNIVERSITIES: RESULTS FROM A SURVEY AMONG ACADEMIC RESEARCHERS[8]

BACKGROUND

There is a widely held concern that various forms of research misbehavior and breaches of research integrity are eroding the public trust in science and weakening the reliability of research results (Steneck 2006). Besides the development of educational materials and guidance documents, research organizations responded with multiple research initiatives that attempt to understand the emergence and causes of research integrity breaches. Academic journals are dominated by discussions regarding prominent misconduct cases (falsification, fabrication, and plagiarism). Similarly, media reports usually focus on the violations of a particular scientist accused of misconduct (Armond et al. 2021). The initial focus of research integrity studies on misconduct and individual behavior has been slowly expanded with studying external, structural, and institutional factors in understanding a wider variety of questionable research practices (Martinson, Anderson, and de Vries 2005, Steneck 2002, 2006). Interests in studying the organizational climate of researchers as external determinants of questionable research practices grow as these could also identify points of interventions for specific research institutions in fostering research integrity (Haven, Tijdink, et al. 2019).

Organizational climate is "*the shared meaning organizational members attach to the events, policies, practices, and procedures they experience and the behaviors they see being rewarded, supported, and expected*." The term "organizational research climate" refers to the current patterns of organizational life and behavior related to research activities among organizational members and leaders (Schein 1999). Organizational research climate studies focus on the institutional work environment of researchers, whether the departmental or university climate strengthens, supports, or erodes research integrity

[8] This chapter is based on an article under review.

(Mumford et al. 2007, Martinson et al. 2016). The Survey of Organizational Research Climate (SOURCE©) instrument has been developed to measure the organizational research integrity climate in academic research settings (Thrush et al. 2007).

The theoretical background of SOURCE© originates in organizational justice as an umbrella term used to refer to individuals' perceptions about the "fairness" of decision-making and resource distribution within organizations and the behavioral consequences of those perceptions (Martinson et al. 2010b). If people perceive injustice in their organization, they are more likely to behave in ways that, in their mind, compensates for the perceived unfairness. In a research climate where perceived injustice is high, researchers would be expected to be more likely to engage in intentional research misconduct (falsification, fabrication, and plagiarism) or questionable research practices (Haven, Tijdink, et al. 2019). The development of SOURCE© has also grounded in the Institute of Medicine and the National Research Council's recommendations in 2002 on open systems conceptual framework for research integrity. It recognizes research integrity as an outcome of processes influenced by multiple factors (institution's visible ethical leadership, socialization and communication processes, and the presence of policies, procedures, structures, and processes to deal with risks to integrity) (Wells et al. 2014).

The SOURCE© instrument was used in several recent studies in various contexts. Wells et al. studied three research intensive universities and found that researchers in different phases of their career perceive the research integrity climate differently (Wells et al. 2014). Doctoral students perceived the climate to be fairer compared to senior scientists in that scholarly integrity was valued. Senior scientists perceived more positively the resources for conducting research responsibly (Wells et al. 2014). This study also found a significant difference regarding SOURCE© scores between different organizational subunits. The scores of some units were twice as negative compared to others or compared to the overall mean scores of the university. This indicates that research climate may vary significantly within institutions. One factor that accounts for these stark differences between subunits was the scientific field.

Another more recent study assessing research integrity climate focused on differences in academic rank and scientific field in two university medical centers and two universities in

Amsterdam (Haven, Tijdink, et al. 2019). The study by Haven et al. also showed differences regarding the research integrity climate perceptions of academic rank and scientific fields. Small fields like the Humanities perceive their department's expectations as more negative compared to other scientific fields. According to this study, the natural sciences perceive the climate more positively. Associate and full professors perceive a more positive research integrity climate than assistant professors, postdocs, and PhD students.

Our study aims to determine how scientists experience the research integrity climate, stratified for academic rank and scientific field, in three large Hungarian universities with doctoral schools covering the full spectrum of disciplines. To our knowledge, this is the first study that investigates research integrity climate in the country.

METHODS

This study was reviewed and approved by the Research Ethics Committees in Hungary (Approval number: IV8159-1/2020/EKU).

Sample

The Doctoral Schools from three universities in Hungary (University of Debrecen, Pécs, and Szeged) were invited to participate. Sixteen Doctoral Schools agreed to participate in the study, with a total of 2557 registered PhD students, postdocs, and professors.

Procedure

The Doctoral Council of each university was contacted to participate and asked to forward the invitation letter for each Doctoral School. The Doctoral Schools that agreed to participate were asked to send by email the letter along with the survey link for every registered researcher (PhD students, postdoctoral trainees, assistant professors, associate professors, and full professors). Two reminders were sent, each one week apart. The data was collected between December 2020 and February 2021.

The survey was created using Qualtrics (Qualtrics, Provo, UT, USA). It consisted of the informed consent and demographic questions (gender, academic rank, and scientific field) and the Survey of Organizational Research Climate - SOURCE©.

Instrument

The Survey of Organizational Research Climate (SOURCE©) is a validated questionnaire developed to assess the organizational climate of research integrity in an academic setting. The questionnaire includes 32 items, 28 items consisting of 7 subscales, two items assessing the global perception of the institutional environment, and two items of global perception of the department or program. The items are scored on a 5-point Likert scale (1 = 'Not at all, 5 = 'Completely'). A sixth option, "No basis for judging" is also offered not to force a response. The survey consists of two parts. The first part contains 11 questions on one's perception of research climate in the wider institution. The second, with 21 items, regards one's perception of research climate in the closer working environment, department, or program.

The 7 subscales assess the following areas: (1) Institutional RCR Resources (6 items). This scale assesses perceptions of the effectiveness of RCR resources and research-related communications, understanding of misconduct reporting procedures, and accessibility of research resources (e.g., policies, experts). (2) Institutional Regulatory Quality (3 items). This scale measures perceptions on how fair and respectful regulatory bodies (e.g., RECs, animal ethics committees) are with researchers and how familiar they are with the research they review. (3) Subunit Integrity Norms (4 items). This scale assesses perceptions on how much the department values scholarly integrity (e.g., honesty, confidentiality). (4) Subunit Integrity Socialization (4 items). This scale assesses perceptions of departmental commitment to socialize junior researchers in RCR effectively. (5) Subunit Advisor /Advisee Relations (3 items). This scale assesses perceptions about fairness, respect, and the availability of advisors to advisees. (6) Subunit Integrity Inhibitors (6 items). This scale measures respondents' perceptions of the department's negative impact on certain conditions (e.g., lack of adequate human or material resources, pressure to publish, competition among researchers). This scale is reversed-scored to express the lack of

inhibitors. (7) Subunit Expectations (2 items). This scale assesses the perceptions related to the fairness of the departmental expectations for publishing and obtaining external funding.

The survey was delivered both in the original version, in English, and in the Hungarian translation. The English version was translated into Hungarian by two different translators. Then, a bilingual researcher familiar with the field of study cross-checked the translation, and the survey was back-translated into English. Any linguistic discrepancies were discussed until a satisfactory version was reached. Respondents could choose between English or the Hungarian version.

Statistical Analysis

All the analyses were performed using IBM SPSS version 24.0. The mean scores for the seven subscales (RCR Resources, Regulatory Quality, Integrity Norms, Integrity Socialization, Advisor/Advisee Relations, Integrity Inhibitors, and Departamental Expectations) were calculated and stratified for academic rank and scientific field. Measurements between groups and within groups were analyzed using the Kruskall-Wallis test, adjusted with Bonferroni correction. The effect sizes data were calculated using eta square, where $\eta^2 = H\text{-}k+1/n\text{-}k$, and the interpretation values are: 0.01- < 0.06 (small effect), 0.06 - < 0.14 (moderate effect) and >= 0.14 (large effect). For the paiwise comparisons, the effect size was calculated using the correlation coefficient employing the formula $r = Z/\sqrt{n}$, and the interpretation values are: 0.1- < 0.3 (small effect), 0.3 - < 0.5 (moderate effect) and >= 0.5 (large effect) (Cohen 2013). Multivariate regression analyses were performed to assess confounding interactions (with gender, age, academic rank, and scientific field) on the variables that were significant.

RESULTS

A total of 2557 people were invited to participate. 758 researchers opened the survey (30%). Of those who opened, 438 completed at least one subscale of SOURCE© (17%). Of the responding researchers, 46% were female, 54% were male, and two researchers did not disclose their gender. 44% are PhD students, 17% are postdocs or assistant professors, and

39% are an associate or full professors. 47% are from the Biomedical Sciences, 23% from Natural Sciences, 18% from Social Sciences, and 12% from Humanities.

Table 1. presents the number of cases, the mean scale score, standard deviation, the number of scales' items, and the reliability coefficient. The Advisor/Advisee Relations scale had the highest score (3.92), Integrity Norms scored slightly lower (3.84), and Subunit Integrity Inhibitors scored the lowest (2.96). RCR Resources, Regulatory Quality, Integrity Socialization, and Departmental Expectations had average scores, ranging from 3.52 and 3.65. The reliability coefficient, expressed by Cronbach's alpha (α), ranged from 0.69 to 0.88.

Table 1. Number of Cases, Mean Score, Standard Deviation, Number of items, and Reliability by scale.

Scales	N	Mean	SD	Items	Cronbach's alpha(α)
RCR Resources	431	3.52	0.93	6	0.883
Regulatory Quality	383	3.65	0.91	3	0.802
Integrity Norms	428	3.84	0.85	4	0.822
Integrity Socialization	430	3.56	0.88	4	0.855
Advisor/ Advisee Relations	432	3.92	0.86	3	0.878
Integrity Inhibitors	430	2.96	0.82	6	0.769
Departmental Expectations	424	3.56	0.96	2	0.693

SOURCE© subscales and academic rank

For every scale, postdocs and assistant professors perceived integrity climate more negatively than PhD students and full or associate professors. In contrast, PhD students perceive more positively than the other groups (Table 2).

Table 2. Scales mean scores and standard deviation across academic rank.

Academic rank	PhD student	Postdoc or assistant professor	Associate or full professor	p*	η^{2**}
RCR Resources	3.68 (.96)	3.28 (.86)	3.43 (.89)	**0.00**	0.03
Regulatory Quality	3.83 (.87)	3.50 (.97)	3.51 (.89)	**0.00**	0.03
Integrity Norms	3.99 (.86)	3.71 (.89)	3.73 (.80)	**0.00**	0.02
Integrity Socialization	3.64 (.95)	3.33 (.85)	3.56 (.80)	**0.01**	0.02
Advisor/Advisee relations	3.99 (.93)	3.69 (.91)	3.93 (.73)	**0.01**	0.02

(Lack of) Inhibitors	3.01 (.88)	2.91 (.80)	2.92 (.77)	0.55	0.00
Expectations	3.76 (.97)	3.19 (.87)	3.51 (.94)	**0.00**	0.05

* Kruskall-Wallis test.

**An effect size of 0.01- < 0.06 is small, 0.06 - < 0.14 is moderate, and >= 0.14 is large.

When analyzed according to the academic rank groups, statistically significant differences were found for six subscales, Institutional RCR resources, Institutional Regulatory quality, Subunit Integrity Norms, Subunit Integrity Socialization, Advisor/Advisee Relations, and Subunit expectations.

Pairwise, adjusted using the Bonferroni correction, Institutional RCR resources scores were higher for PhD students (mean score= 3.68, 95% CI 3.54-3.81) when compared with postdocs and assistant professors (mean score= 3.28, 95% CI 3.08-3.48, p=0.001, r=0.22), and to associate or full professors (mean score= 3.43, 95% CI 3.30-3.57, p=0.014, r=0.15). There were no significant differences between the other pairs. In the Subunit Regulatory quality, PhD students scored higher (mean score= 3.83, 95% CI 3.70-3.97) when compared with postdocs and assistant professors (mean score= 3.50, 95% CI 3.25-3.74, p=0.027, r=0.17), and to associate or full professors (mean score= 3.51, 95% CI 3.37-3.66, p=0.002, r=0.19). In Integrity Norms subscale, PhD students scored higher (mean score= 3.99, 95% CI 3.87-4.11) when compared to postdocs and assistant professors (mean score= 3.71, 95% CI 3.50-3.92, p=0.032, r=0.16), and to associate or full professors (mean score= 3.73, 95% CI 3.61-3.86, p=0.006, r=0.16). There was no evidence of a difference between the other pairs. The PhD students' scores (mean score= 3.64, 95% CI 3.50-3.77) in the Subunit Integrity Socialization were higher than postdocs or assistant professors (mean score= 3.33, 95% CI 3.14-3.53, p=0.008, r=0.19). The PhD students' scores (mean score= 3.99, 95% CI 3.86-4.12) were also higher in the Advisor /Advisee Relations subscale than postdocs or assistant professors (mean score= 3.69, 95% CI 3.48-3.90, p=0.007, r=0.19). Subunit expectations scores were the lowest for postdocs or assistant professors (mean score= 3.19, 95% CI 2.99-3.39) when compared with highest, for PhD students (mean score= 3.76, 95% CI 3.62-3.90, p<0.001, r=0.28), and with associate or full professors (mean score= 3.51, 95% CI 3.37-3.66, p=0.040, r=0.16). PhD students also scored higher when compared with an associate or full professor (p=0.032, r=0.14). Regulatory quality, Integrity Norms, and

Advisor /Advisee Relations were confounded by the scientific field. However, when corrected for confounding with the scientific field, the associations remained significant (see Table 3).

Table 3. Regression models of SOURCE subscales and academic rank.

Academic rank	Postdoc		Professor		PhD student	
Subscale *F (p,df)*	Beta	(CI)	Beta	(CI)	Ref	Ref
RCR resources 6.188 (.002, 2)	-.161	(-.644,-.148)	-.123	(-.431,-.040)	-	-
Regulatory Quality 6.177 (.002,2)	-.143	(-.614,-.091)	-.180	(-.532,-.131)	-	-
Integrity Norms 5.193 (.006,2)	-.116	(-.491,-.033)	-.128	(-.404,-.045)	-	-
Integrity Socialization 3.227 (.041, 2)	-.121	(-.519,-.043)	-.022	(-.228,.147)	-	-
Advisor/Advisee relations 3.432 (.003,2)	-.129	(-.523,-.063)	-.014	(-.206,.155)	-	-
Expectations 9.941 (<.001,2)	-.221	(-.819,-.304)	-.120	(-.439,.034)	-	-

F-tests (*F*) and the associated p-value and degrees of freedom. Regression coefficients (Beta) and confidence intervals (CI) adjusted for confounding factor (scientific field).

SOURCE© subscales and scientific field

Overall, Natural sciences perceived integrity climate more negatively than the other scientific fields, and Humanities more positively (Table 4).

Table 4. Scales mean scores and standard deviation across scientific fields.

Scientific field	Biomedical sciences	Natural sciences	Social sciences	Humanities	p*	η^2**
RCR Resources	3.52 (.97)	3.42 (.89)	3.56 (.96)	3.74 (.84)	0.18	0.00
Regulatory Quality	3.79 (.85)	3.49 (.87)	3.62 (1.06)	3.80 (.87)	**0.02**	0.02
Integrity Norms	3.79 (.94)	3.75 (.81)	3.98 (.79)	4.10 (.75)	**0.02**	0.02
Integrity Socialization	3.53 (.95)	3.47 (.80)	3.69 (.88)	3.69 (.94)	0.11	0.01
Advisor/Advisee relations	3.80 (.94)	3.89 (.81)	3.99 (.85)	4.20 (.71)	**0.04**	0.01
(Lack of) Inhibitors	3.09 (.85)	2.87 (.76)	2.81 (.88)	3.12 (.83)	**0.01**	0.02
Expectations	3.58 (.90)	3.46 (.93)	3.64 (1.10)	3.73 (.95)	0.18	0.00

* Kruskall-Wallis test.

**An effect size of 0.01- < 0.06 is small, 0.06 - < 0.14 is moderate, and >= 0.14 is large.

The analysis of scientific fields provided statistically significant differences between scientific fields in four subscales: Regulatory quality, Integrity Norms, Advisor /Advisee Relations, and Inhibitors.

In pairwise analyses, Biomedical sciences (mean score= 3.78, 95% CI 3.64-3.93) scored significantly higher on Regulatory quality than Natural Sciences (mean score= 3.49, 95% CI 3.35-3.64, p=0.025, r=0.17). Natural sciences score significantly lower (mean score= 3.75, 95% CI 3.62-3.87) on Integrity Norms than Humanities (mean score= 4.10, 95% CI 3.88-4.31, p=0.032, r=0.19). Humanities (mean score= 4.20, 95% CI 4.00-4.40) also scored significantly higher than Biomedical sciences (mean score= 3.80, 95% CI 3.64-3.96) on Advisor /Advisee Relations (p=0.046, r=0.19). There was no evidence of a difference between the other pairs. When corrected for confounding with the academic rank, the associations remained significant (see Table 5).

Table 5. Regression models of SOURCE subscales and scientific field.

Scientific field	Natural Sciences		Social Sciences		Humanities		Biomedical Sciences	
Subscale *F (p,df)*	Beta	(CI)	Beta	(CI)	Beta	(CI)	Beta	(CI)
Regulatory Quality 2.772 (.041,3)	-.153	(-.500,-.075)	-.110	(-.517,.006)	-.006	(-.334,.300)	Ref	-
Integrity Norms 3.013 (.030,3)	-.186	(-.592,-.058)	-.070	(-.455,.147)	Ref	-	-.156	(-.558,-.011)
Advisor/Advise e relations 2.946 (.033,3)	-.172	(-.571,-.035)	-.103	(-.528,.073)	Ref	-	-.219	(-.678,-.129)

F-tests (*F*) and the associated p-value and degrees of freedom. Regression coefficients (Beta) and confidence intervals (CI) adjusted for confounding factor (academic rank).

DISCUSSION

This study aimed to assess Hungary's integrity climate and the differences across scientific fields and academic rank using the SOURCE© questionnaire. To our knowledge, this is the first study to investigate the research climate in Hungarian institutions.

Overall, the study results highlight some critical points of the integrity climate in Hungary. The study findings show that overall mean scores for all scales were below 4.00. The results contrast with the findings from the U.S. by Wells et al. (2014) and Martinson et al. (2016) and are similar to the results from Amsterdam by Haven et al. (2019). The scales Advisor/Advisee relations and Integrity norms had the highest scores in the study, while (the lack of) Integrity Inhibitors scored the lowest. Thus, although study participants

perceive more positively the attitude of supervisors and the support of norms, they are exposed to different factors of Integrity Inhibitors. The integrity inhibitors express the presence of certain conditions, such as lack of resources (human or material), pressure to obtain funding, pressure to publish, suspicion, and competition among researchers.

Publication pressure and competition among researchers can be harmful to the integrity of science. These factors affect the willingness to open sharing of information, interfere with the peer-review process, and lead to detrimental research practices (Anderson et al. 2007, Fanelli 2010). The Hungarian findings had the worst scores on this scale, contrasting to the findings from The Netherlands (Haven, Tijdink, et al. 2019) and the U.S. (Wells et al. 2014). This might be explained by similarities in competitiveness and pressure to publish trends in academia that are not accompanied by similar human and material resources at these research institutions. Researchers at the studied universities often have to balance their research activities with teaching duties. Moreover, Hungarian universities have much more modest financial resources for scientific research than universities in the U.S. or the Netherlands.

Academic Rank

Academic rank has shown to be significantly related to six scales. Surprisingly, the mean patterns by ranks were identical across all scales. Overall, PhD students perceive more positively than the other groups, while postdocs and assistant professors perceive more negatively. These results are similar to those found by Wells et al. (2014) in the U.S., and contrast to the results found by Haven et al. (2019) in Amsterdam, where PhD students scored lower than professors across all scales. The potential explanations by Wells et al. (2014) for the low scores of postdocs and assistant professors were shorter time in the research environment and less contact to the institution. Recent surveys (Woolston 2020b, a, Afonja et al. 2021) show that postdocs are exposed to higher pressures to publish and obtain funding than PhD students due to an unstable work position and a competitive environment. The little job security exposes them to higher expectations to secure their position. The surveys have also shown that postdocs in their academic position get little or no guidance or resources for their work. These findings could explain the lowest scores of postdocs and assistant professors.

Scientific field

The scientific field was also significantly related to perceived integrity climate in four scales. However, the mean patterns by disciplines are not similar across all scales, except Humanities, which had the highest scores for every scale. The Humanities results contrast with the results found by Haven et al. (2019), where Humanities scored lowest in most of the scales. The high scores of the Humanities could be explained by their separateness from the research system of the more empirical and quantitative scientific fields, especially fields like biomedical research. In Hungarian universities, Humanities still focus on books as the major research output. International competitiveness is less present, and some of the research integrity's challenges are absent from their everyday work. In Regulatory quality, pairwise differences were found between Biomedical and Natural sciences. A higher score for Biomedical sciences is expected as regulatory bodies such as research and animal ethics committees are essential to biomedical research. The same might not apply to natural sciences disciplines, such as Mathematics or Physical sciences.

In the Integrity norms scale, researchers from Humanities perceive how much the department values scholarly integrity more positively than those from Natural Sciences. In Advisor/Advisee Relations, Humanities scored higher than Biomedical sciences. Recent evidence highlights ethical issues involving mentoring in Biomedical Sciences, such as misalignment of goals, poor communication, and failure to acknowledge advisee's contribution (Kow et al. 2020). However, little is known about this context in Humanities. One potential explanation could refer to the differences regarding the disciplinary traditions of mentoring in PhD student-Supervisor or Postdoc trainee-Supervisor relationships in the Humanities.

Limitations

Our study has some limitations that should be addressed. First, there was low interest in participation by Doctoral Schools. Only 24% of the invited Doctoral Schools actively agreed to participate by sending us feedback. This might demonstrate a low interest of Doctoral Schools to investigate institutional integrity climate. Although the study was completely anonymous and there was no comparison between units or departments, fear of

ranking or retaliation is a potential explanation, as the results can be sensitive. Consequently, there is a possibility that the Doctoral Schools that have agreed to participate are the ones that most foster integrity, which creates a selection bias.

Second, as there was no data collection regarding departments or programs, it was not possible to investigate differences between sub-units. In the results by Wells et al. (2014), differences across small organizational units account for a great part of the variability in the scales. Third, the response (30%) and completion (17%) rates were relatively low. However, our rate is comparable to that of other web-based surveys (Haven, Tijdink, et al. 2019, Cook, Heath, and Thompson 2000).

CONCLUSION

Research integrity climate is a strong factor that influences an individual's behavior. Negative perceptions of the research climate were associated with a higher likelihood of engaging in detrimental research practices (Crain, Martinson, and Thrush 2013). Hence, a strong research integrity culture can lead to better research practices and responsible conduct of research (Forsberg et al. 2018). While everyone involved in the research endeavor is responsible for the ethics and integrity of science, there has been a call for research performing organizations to provide measures to strengthen integrity (Forsberg et al. 2018, Mejlgaard et al. 2020).

Institutional initiatives for promoting and fostering integrity should be evidence-supported and as tailored as possible (Viđak et al. 2021). Organizational climate investigations are a valuable instrument to identify the strengths and weaknesses of each subgroup to develop targeted initiatives and institutional policies. Our study indicated some critical points regarding the integrity climate in Hungarian universities.

If we look at scientific fields, overall, Natural sciences perceived integrity climate more negatively than other fields. Therefore, department leaders should develop initiatives to address the weaknesses and to foster better integrity climates. According to academic ranks, postdocs and assistant professors perceived the integrity climate more negatively on every scale. The results suggest that institutions should pay more attention to early career

researchers, especially those in insecure and transitory work positions. They should provide RCR resources, socialize them in RCR, and set more reasonable expectations.

The low scores on integrity inhibitors indicate that responsible research conduct is impacted by non-ideal research conditions related to resources, pressure, and interpersonal relations. These factors require further detailed studies developed to these institutional contexts.

CHAPTER 5 – PERCEPTIONS OF PUBLICATION PRESSURE: SURVEY WITH RESEARCHERS FROM HUNGARIAN UNIVERSITIES[9]

BACKGROUND

Scientific publications are a vital part of scientific communication and progress. However, publishing has become a necessity in a competitive context where scientific output is used as a metric for promotion, funding, and tenure. Researchers are under pressure for publishing and have to balance teaching or other academic activities with scientific production in most cases (Haven, Bouter, et al. 2019).

Publication pressure can be understood as the perceived culture where researchers feel they have to publish (Haven, de Goede, et al. 2019). This feeling can be both external and self-inflicted. In a focus group study on publication pressure (Tijdink, Schipper, et al. 2016), participants reported that they feel their positions in academia depend on publishing articles in high-impact journals and obtaining funding. They have also reported that they feel being judged based on their number of publications.

Publication pressure and competitiveness can motivate scientific productivity and advancements (Fanelli, Costas, and Lariviere 2015). On the other hand, many studies have associated publication pressure with detrimental aspects for individual researchers and the whole science ecosystem. A recent systematic review found that publication pressure can be the main reason for choosing a predatory journal deliberately (Mills and Inouye 2021). Furthermore, this association was also found with publication bias (Grimes, Bauch, and Ioannidis 2018), redundant publications (Yank and Barnes 2003), problems in reproducibility (Baker 2016), and low quality in other academic roles (Copeland 2022).

Whether publication pressure can be associated or not with the likelihood of scientific misconduct remains unclear. No clear support was found in the association of duplicate images (Fanelli et al. 2018) or retracted papers (Fanelli, Costas, and Lariviere 2015) and publication pressure. Differently, publication pressure was reported as a reason for

[9] This chapter is based on an article under review.

scientific misconduct in survey studies (Shamsoddin et al. 2021, Maggio et al. 2019, Paruzel-Czachura, Baran, and Spendel).

The reward systems in academia are significant contributors to the "publish or perish" culture. Therefore, the country, academic position, or even departments may perceive publication pressure differently. Many studies have investigated publication pressure in the US (Fanelli 2010) and western European countries (Haven, Bouter, et al. 2019, Tijdink, Schipper, et al. 2016). To our knowledge, we are the first to assess perceived publication pressure and its relationship with academic rank, disciplinary field, and gender in an Eastern European country. Therefore, this study aims to assess perceived publication pressure and its relationship with academic rank, scientific field, and gender in Hungarian universities.

METHODS

This study was reviewed and approved by the Research Ethics Committees in Hungary (Approval number: IV8159-1/2020/EKU).

Sample

The Doctoral Schools from three universities in Hungary (University of Debrecen, Pécs, and Szeged) were invited to participate. Sixteen Doctoral Schools in Hungary agreed to participate in the study, with 2557 registered PhD students, postdocs, and professors.

Procedure

We contacted the Doctoral Council of each university and asked them to forward the invitation letter for each Doctoral School. The Doctoral Schools were then asked to send by email the letter along with the survey link for every registered researcher (PhD students, postdoctoral trainees, assistant professors, associate professors, and full professors). Two reminders were sent, each one week apart.

The survey was created using Qualtrics (Qualtrics, Provo, UT, USA). The survey started with the informed consent and demographic questions (gender, academic rank, and

disciplinary field). The second part included the revised Publication Pressure Questionnaire and the Survey of Organizational Research Climate.

Instrument

The publication pressure perception was assessed using the revised Publication Pressure Questionnaire (PPQr) (Haven, de Goede, et al. 2019). The PPQr was translated into Hungarian and also made available for the respondents in English. The translation of the questionnaire was conducted in four steps: (1) The original English version was translated by two different translators, emphasizing conceptual rather than literal translations and the need to use a natural and acceptable language for the most public audience; (2) A bilingual researcher, familiar with the terminology of the area covered by the instrument, together with the two translators, met to resolve any discrepancies; (3) The questionnaire in Hungarian was back-translated to English, and any linguistic discrepancies were discussed until a satisfactory version was reached; (4) The final Hungarian version was finalized.

The PPQr is a reliable and validated instrument and includes three subscales, Publication Stress, Publication Attitude, and Publication Resources. The Publication Stress subscale assesses the perceived stress resulting from the pressure a researcher feels to publish. The Publication Attitude assesses researchers' attitudes regarding the publication pressure. The Publication Resources includes inhibitor factors to publication pressure. The questionnaire consists of 18 items, and the items are scored on a 5-point Likert scale (1 = 'totally disagree', 5 = 'totally agree'). Positive items such as "I feel sufficiently capable of writing an academic paper" were inversely scored, so higher scores express high perceived publication stress, negative attitude towards publication climate, and lack of resources.

Statistical Analysis

All the analyses were performed using IBM SPSS version 24.0. The mean scores for the three subscales (Stress, Attitude, and Resources) were calculated and stratified for academic rank, disciplinary field, and gender. Measurements between groups and within groups were analyzed using the Kruskal-Wallis test, adjusted with Bonferroni correction. The effect sizes data were calculated using eta square, where $\eta^2 = H\text{-}k+1/n\text{-}k$, and the interpretation values are: 0.01- < 0.06 (small effect), 0.06 - < 0.14 (moderate effect) and $>=$

0.14 (large effect). For the paiwise comparisons, the effect size was calculated using the correlation coefficient employing the formula $r = Z/\sqrt{n}$, and the interpretation values are: 0.1- < 0.3 (small effect), 0.3 - < 0.5 (moderate effect) and >= 0.5 (large effect) (Cohen 2013). Multivariate regression analyses were performed to assess confounding interactions (with gender, age, academic rank, and scientific field) on the variables that were significant.

RESULTS

A total of 2557 people were invited to participate. 758 researchers opened the survey (30%). Of those who opened, 408 completed at least one subscale of the survey (16%). Of the responding researchers, 46% were female, 54% were male, and two researchers did not disclose their gender. 45% are PhD students, 17% are postdocs or assistant professors, and 38% are associate or full professors. 31% are from the Biomedical Sciences, 39% from Natural Sciences, 18% from Social Sciences, and 12% from Humanities.

Overall, the investigated researchers score highest for Stress (M = 3.06), followed by Attitude (M = 3.03) and Resources (M = 2.30). The reliability coefficient, expressed by Cronbach's alpha (α), ranged from 0.73 to 0.82.

Publication pressure and scientific field

There were no statistically significant differences between the publication pressure subscales and the disciplinary fields (Table 1). Figure 1. presents the means and the error bars express 95% confidence intervals.

Table 1. PPQ Scales across scientific fields.

Scales	Biomedical sciences	Natural sciences	Social sciences	Humanities	p*	η^2**
Stress	2.89 (.94)	3.12 (.92)	3.23 (.86)	3.09 (.95)	.077	0.01
Attitude	2.95 (.78)	3.08 (.74)	3.14 (.75)	2.84 (.79)	.056	0.01
Resources	2.28 (.68)	2.25 (.72)	2.43 (.70)	2.29 (.70)	.430	0.00

* Kruskal-Wallis test.

**An effect size of 0.01- < 0.06 is small, 0.06 - < 0.14 is moderate, and >= 0.14 is large.

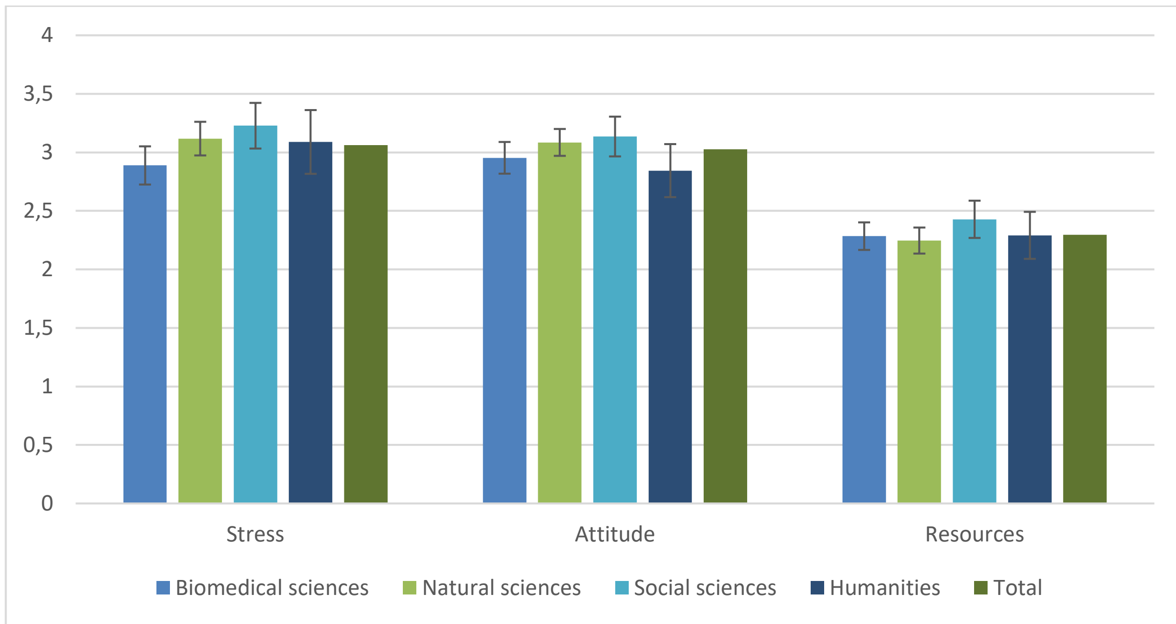


Figure 1. PPQ Scales means and 95% CI error bars across scientific fields in Hungary.

Publication pressure and academic rank

There was a statistically significant difference between groups in the Stress subscale as determined by the Kruskal-Wallis test (Table 2). However, the adjusted pairwise analysis was not statistically significant. In Resources, both PhD students (mean score= 2.59, 95% CI 2.49-2.68, p<.001, r=0.18), postdocs and assistant professors (mean score= 2.25, 95% CI 2.09-2.40, p=.012, r=0.19) perceive significantly less Resources than associate professors (mean score= 1.97, 95% CI 1.89-2.07), even after corrected for confounding factors (gender and age) (Table 3). PhD students also perceive less Resources than postdoc and assistant professors (p=.002, r=0.51). Figure 2. presents the means and 95% confidence intervals.

Table 2. PPQ Scales and academic rank.

Scales	PhD student	Postdoc or assistant professor	Associate or full professor	p*	η^2**
Stress	3.13 (.83)	3.20 (1.02)	2.92 (.98)	**.033**	0.01
Attitude	3.06 (.74)	3.11 (.78)	2.95 (.78)	.133	0.01
Resources	2.59(.68)	2.25 (.67)	1.97 (.59)	**.000**	0.16

* Kruskal-Wallis test.
**An effect size of 0.01- < 0.06 is small, 0.06 - < 0.14 is moderate, and >= 0.14 is large.

Table 3. Regression models of PPQr subscales and academic rank.

Academic rank	Postdoc		Professor		PhD student	
Subscale *F (p,df)*	Beta	(CI)	Beta	(CI)	Ref	Ref
Stress 3.237 (.040, 2)	.691	(-.182,-.380)	-.012	(-.345,-.301)	-	-
Resources 38.323 (<.001,2)	-.164	(-.505,-.104)	-.385	(-.786,-.324)	-	-

F-tests (*F*) and the associated p-value and degrees of freedom. Regression coefficients (Beta) and confidence intervals (CI) adjusted for confounding factors (age, gender, and academic rank).

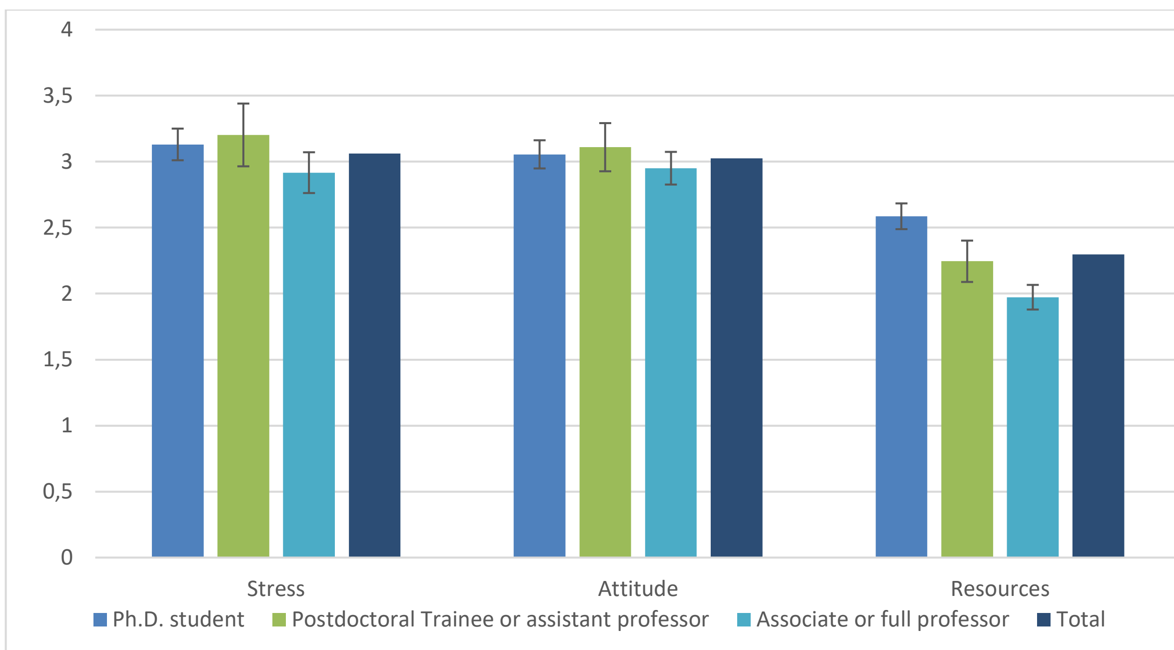


Figure 2. PPQ Scales mean and 95% CI error bars across academic ranks.

Publication pressure and gender

Gender differences were found in two scales, Stress and Resources (Table 4). In the Stress scale, female researchers (mean score= 3.28, 95% CI 3.15-3.40) scored significantly higher than male researchers (mean score= 2.89, 95% CI 2.76-3.00, p<.001, r=0.22), even after correction for confounding factors (age and academic rank). In the Resources scale, female researchers also scored significantly higher than male researchers, but the difference was not significant after correction for confounding (academic rank).

Table 4. PPQr Scales and gender.

Scales	Female	Male	p*	η²**
Stress	3.28 (.89)	2.89 (.91)	**.000**	0.04
Attitude	3.03 (.79)	3.01 (.74)	.853	0.00
Resources	2.39 (.73)	2.21 (.66)	**.028**	0.01

* Kruskal-Wallis test.

**An effect size of 0.01- < 0.06 is small, 0.06 - < 0.14 is moderate, and >= 0.14 is large.

Table 5. Regression models of PPQr subscales and gender.

Gender	Female		Non disclosed	
Subscale *F (p,df)*	Beta	(CI)	Beta	(CI)
Stress 9.728 (<.001,2)	.183	(.154,.525)	-.024	(-.952,1.599)
Resources 3.503 (.031,2)	.073	(-.030,.235)	-.0.39	(-.523,1.296)

F-tests (*F*) and the associated p-value and degrees of freedom. Regression coefficients (Beta) and confidence intervals (CI) adjusted for confounding factors (age and academic rank).

DISCUSSION

Quantitative indicators of research output have been traditionally used to determine parameters to obtain research grants, promotions, and prestige. While competition and pressure to publish can be positively indicated as part of the academic endeavor and science growth, it should not be done at the expense of personal burden and research excellence.

The researcher's assessment practices and the way publication outputs are evaluated differ across countries. In Hungary, there is no established or standardized framework to assess researcher's publication outputs in academic working environments. Universities rarely assess their academic staff on a yearly basis, but typically apply a variety of research assessment practices in the case of promotion decisions (PhD degree, Habilitation procedures or professorships). In the field of biomedical and life sciences, the use of scientometrics, especially the role of impact factor, and the quantitative use of IF and other citation metrics is a driving force of research assessment (Benedictus, Miedema, and Ferguson 2016, Curry 2018). In the highly centralized grant system of the country, the main research funding organization also emphasizes these "traditional" scientometrics in

defining scientific excellence when reviewing grant applications. There were some scattered discussions regarding the limits and biases of the use of these scientometrics in defining academic careers (Seglen 1997), but new initiatives for research assessment did not appear within the major research funding or performing organizations of the country.

Pressure to publish is often mentioned as the motivational factor in misconduct cases and questionable research practices (Davis, Riske-Morris, and Diaz 2007a). Studies have also found personal negative effects, such as burnout syndrome. Therefore, our study aimed to investigate publication pressure factors in Hungarian universities and their differences according to the career stage, scientific field, and gender.

Our findings show no differences in perceptions of publication pressure between scientific fields. The study by Haven et al. (2019) found small differences across scientific fields in Stress and Resources scales. However, the differences were greater when associated with the academic rank.

In the academic rank analysis, our results show that PhD students perceive a greater lack of resources than postdocs and professors. Likewise, postdocs and assistant professors also perceive a greater lack of resources than associate and full professors. This result is expected, as senior researchers have more experience to cope with publication procedures and pressure. However, in the face of these results, senior researchers should better assist junior researchers during the publication process and improve academic mentoring.

The results also show that female researchers perceive higher stress than male researchers, regardless of the academic rank or scientific field. Gender differences in academia have been broadly documented from the perspectives of citations, publications, salary, recognition, and promotions (Holman, Stuart-Fox, and Hauser 2018, Larivière et al. 2013, Shen 2013). Women are more likely to get less funding (Wadman 2019), work support (Else 2019), and be hit harder in the work-life imbalance. Moreover, a recently published study by Huang et al. (2020) shows that career length can largely explain gender gaps in science. Women dropout rates in academia are higher at every career stage, subsequently affecting their productivity and increasing indicator disparities. The consequences of a gender-biased context on publication pressure may collaborate to a higher dropout rate and

perpetuate gender inequality in academia. Accordingly, it is essential that organizational structures develop policies or strategies to address gender gaps in academia.

Our study has some limitations that should be addressed. First, although similar to other surveys (ref), our response rate was relatively low in Hungary (16%). One potential consequence is the underestimation of our results, as the invited researchers might not have answered because they were busy working on their publications. On the other hand, our data were collected in late 2020 and early 2021 amidst the COVID-19 pandemic. Evidence shows that some research fields have come under greater pressure for publication, in addition to the consequences of the home office, such as the balance of domestic tasks and work (Collins 2020). This scenario could have led to an overestimation of our results.

CONCLUSION

Our findings show that pressure to publish similarly affects all scientific fields. However, PhD students and female researchers are affected differently. PhD students perceive a greater lack of resources compared to the other groups, and women researchers perceive a greater stress. Our study highlights the need to improve mentoring during early-career researchers' development. It also emphasizes the importance of organizational structures developing policies or strategies to address gender differences in academia.

CHAPTER 6 – DISSERTATION DISCUSSION

This dissertation mainly focused on the Hungarian context and its characterization regarding research integrity. In this discussion chapter, I summarize the main findings, address some limitations of this work, suggest points for future research, and provide some recommendations for fostering research integrity.

Main findings

Chapter 2 analyzed and discussed how research integrity and ethics cases are represented in the scientific literature. The results show that scientific literature is dominated by discussions regarding prominent cases, primarily involving fabrication and falsification. Moreover, there is an overrepresentation of biomedical research cases. It is possible to conclude that the representation of research integrity and ethics cases in the literature might diverge the attention of the academic community from relevant but less visible violations, such as questionable research practices. **Chapter 3** characterized the PhD students in Hungary and assessed disciplinary differences concerning understandings and attitudes of responsible conduct of research. The results show that PhD students in Hungary believe that they better understand RCR standards than they do in practice. In this study, authorship assignment and data handling are the most causes of uncertainties. Although there are known disciplinary differences, the results generally showed differences only in authorship, and data practices. Moreover, the findings also show a high prevalence of admitted and observed QRPs. **Chapters 4** and **5** assessed organizational and structural factors that influence a researcher's behavior regarding RCR. **Chapter 4** assessed the perceptions on integrity climate and its relationship with academic rank and scientific field in three universities in Hungary. The results show that postdocs and assistant professors perceived integrity climate more negatively than PhD students and full professors in every survey scale. Institutions should pay more attention to early career researchers, especially those in insecure and transitory work positions. They should provide RCR resources, socialize them in RCR, and set more reasonable expectations. Furthermore, there are disciplinary differences that should be considered when developing initiatives to foster better integrity climates. **Chapter 5** assessed perceived publication pressure and its relationship with

academic rank, gender, and scientific field in Hungary. The findings show that there are no disciplinary differences regarding the perceived publication pressure. However, early-career and women researchers are more exposed to publication pressure.

Strengths and limitations

This dissertation provided a better understanding and evidence of research integrity, especially in Hungary. Despite being part of several European initiatives, there is little evidence about research integrity in the Hungarian context. We looked globally into research integrity and ethics cases and Hungary's individual, institutional, and structural factors.

However, the dissertation has some limitations that deserve our attention. First, we have used only quantitative methods in our analyses. Although qualitative methods, through focus group sessions, were initially planned, the COVID-19 pandemic restrained our contacts and delayed our data collection, preventing their submission for this dissertation. We believe that the combination of qualitative and quantitative methods would have added value to the discussion and provided a more robust evidence base for future recommendations.

The other point that deserves our attention is the response bias. In the three conducted surveys, we had low response rates, although similar to other web-surveys (Cook, Heath, and Thompson 2000). The low number of responses does not imply statistical limitations. However, in these cases, the group of respondents could potentially differ from the non-respondents. It is hard to predict the reason for the low response rate. The studies were conducted amid the COVID-19 pandemic, where academics often had to balance their responsibilities and domestic activities. Therefore, the groups' decision to participate in the surveys might have been different. We tried to compare the demographic data of the respondents and our population. However, they were similar.

Another limitation regards the used instruments for the surveys. The SOURCE© and the Revised Publication Pressure Questionnaire (PPQr) were instruments validated in the Hungarian language, but the same does not apply to the INTEGRITY survey. The

INTEGRITY survey was developed by a European-funded project, based on interviews. The lack of validation in the Hungarian language could have potentially affected the survey's understanding.

Suggestions for future research

We are currently assessing the RI guidelines and safeguards in Hungary, and we hope to make the results available soon. We have conducted the same study in Brazil (Armond and Kakuk 2021), and the results provided some important insights on the institutional policies differences across universities, and the importance of institutionalizing RI. For future research, we would like to suggest two main approaches. First, qualitative studies with different stakeholders are essential to provide robust evidence that can be translated into policies. Focus groups with all the major players involved in the research endeavor could provide important insights that the inflexibility of quantitative methods cannot. Systematic data collection based on in-depth interviews with key informants, representatives of RPOs, RFOs, Scientific Societies could also provide important insights.

Second, different interventions to foster research integrity have been studied (Viđak et al. 2021, Katsarov et al. 2021, Sørensen et al. 2021, Marusic et al. 2016). However, most of these interventions are context-dependent. Therefore, in light of the importance of assessing potentially valuable interventions, we highly suggest studies that evaluate the efficacy and effectiveness of specific research integrity interventions within the Hungarian research environment.

Practical implications and recommendations

The public discourse on research integrity mainly evokes prominent misconduct cases (Armond et al. 2021). The cases, such as the Hwang case (Kakuk 2009) or Stapel case (Stroebe, Postmes, and Spears 2012), triggered important debates on research integrity and reliability. But most of these debates focused on the individualization of the cases, understanding motivations, and personal behavior. Although personal traits, such as Machiavellianism, Narcissism, self-esteem, and psychopathy, may affect the likelihood of

research misbehavior (Tijdink, Bouter, et al. 2016), these cases are probably rarer and more challenging to deal with.

Throughout this dissertation, we advocated for an effective promotion of research integrity. But who should take the steps towards it? Who is responsible for the interventions? And what are the best interventions to foster integrity? This dissertation does not cover all the aspects, but the following section combine our findings and the current practices to provide some recommendations.

Research performing organizations and departmental leaders

RPOs are key in fostering research integrity. Their role, however, does not include only damage control, such as investigation and punishment of misbehavior cases. This simplistic perspective ignores the institutional responsibilities and their position as part of the problem (Forsberg et al. 2018). Lack of appropriate training, guidance, and policies, perversive incentives, and the quest for quantity over quality are some institutional factors that can undermine research integrity. Integrity culture must be internalized, learned by practice and role models. On that account, RPOs should provide conditions to strengthen research integrity culture.

Based on our findings, it is important for RPOs and department leaders to diagnose their weaknesses and strengths across departments to provide effective training on RCR tailored to their needs and contexts. Evidence shows that faculty members themselves are in doubt about the appropriate standards of RCR (Alfredo and Hart 2011, Martinson, Anderson, and de Vries 2005). Therefore, the training should not be applied only for early-career researchers but also for supervisors, exploring the gray areas and the changing research environment.

It is also essential to provide codes and guidelines about research integrity (Armond and Kakuk 2021). However, the regulations do not stand alone. They should be practiced and debated in a safe and open environment. RPOs should develop permanent mechanisms to deal with misbehavior, such as a research integrity committee, with safe whistle-blowing channels and transparency when dealing with misconduct cases. Moreover, institutions should rethink their criteria for researcher assessments in hirings and promotions, giving

less attention to narrowly defined individual metrics. It is also necessary to value other activities beyond research, such as teaching, administrative, and mentoring activities (Copeland 2022). Different approaches were formulated to provide a fairer researcher assessment criteria, such as the Hong Kong Principles (Moher et al. 2020), San Francisco Declaration on Research Assessment (DORA), and other national initiatives (Hatch and Curry 2020, Benedictus, Miedema, and Ferguson 2016). These approaches propose improving the way scientific outputs are evaluated, tailoring for the institutional needs.

Supervisors and mentors

The responsibilities of supervisors, advisors, and mentors are different, although the terms are often used interchangeably. All of them are important for an academic or a researcher's development, and they should act as a role model (Bird 2001). However, while research supervisors/advisors are assigned to help an advisee fulfill the requirements with the university, mentors are those willing to share the knowledge and experiences to help in the professional development of a mentee.

RCR is mainly learned from practice by upholding high standards and setting a good example. Being a role model is essential, but it is not enough. Therefore, supervisors and advisors must embrace the mentoring role, keeping in mind the power imbalances (Alfredo and Hart 2011). For that, we suggest that they keep an open dialogue of RCR with their mentees, communicate clearly their expectations regarding best practices and publications, explain why and when some practices are acceptable, and monitor their research activities, such as data collection and analysis.

Early-career researchers

The responsibility of fostering research integrity bears on everyone involved in the research endeavor. Thus, we suggest early-career researchers take a proactive step towards it. They should learn about the policies and regulations relevant to their field and maintain high standards in the research performance and dissemination. Furthermore, it is important to keep an honest and open dialogue with their supervisors. Communicate about their research

progress, including flaws, doubts, and negative results. Communicate also clearly about their expectations regarding their career development and research outputs.

Concluding remarks

As mentioned many times in this dissertation, fostering research integrity requires combined efforts from all the stakeholders. Research funding organizations, scientific journals, and policymakers should also be accounted for, although they have not been described extensively here.

Changing an individual habit could be difficult, but changing a whole culture is even more challenging. The research system is dynamic and involves the relationship of numerous different actors. Thus, promoting research integrity requires frequent reflection on the subject and an awareness that this discussion should be permanent.

SUMMARY

Misconduct cases and questionable research practices (QRPs) have questioned the reliability and objectivity of science and undermined public trust in science. To curb research misbehavior and foster research integrity, it is necessary to understand the research system, its players, and incentives. Therefore, the main objective of this dissertation was to provide some evidence on research integrity and its differences across disciplines, most specifically to the Hungarian contexts, to drive future initiatives to foster research integrity in the country.

A scoping review and three survey-based studies were conducted. **Chapter 2** analyzed and discussed how research integrity and ethics cases are represented in the scientific literature. The results show that scientific literature is dominated by discussions regarding prominent cases, primarily involving fabrication and falsification. Moreover, there is an overrepresentation of biomedical research cases. It is possible to conclude that the representation of research integrity and ethics cases in the literature might diverge the attention of the academic community from relevant but less visible violations, such as questionable research practices. **Chapter 3** characterized the PhD students in Hungary and assessed disciplinary differences concerning understandings and attitudes of responsible conduct of research. The results show that PhD students in Hungary believe that they better understand RCR standards than they do in practice. In this study, authorship assignment and data handling are the most causes of uncertainties. Although there are known disciplinary differences, the results generally showed differences only in authorship practices. Moreover, the findings also show a high prevalence of admitted and observed QRPs. **Chapters 4** and **5** assessed organizational and structural factors that influence a researcher's behavior regarding RCR. **Chapter 4** assessed the perceptions on integrity climate and its relationship with academic rank and scientific field in three universities in Hungary. The results show that postdocs and assistant professors perceived integrity climate more negatively than PhD students and full professors in every survey scale. Furthermore, there are disciplinary differences that should be considered when developing initiatives to foster better integrity climates. **Chapter 5** assessed perceived publication pressure and its relationship with academic rank, gender, and scientific field in Hungary. The findings show that there are no disciplinary differences regarding the perceived publication pressure. However, early-career and women researchers are more exposed to publication pressure.

This dissertation provided a better understanding and evidence of research integrity, especially in the Hungarian context. In **Chapter 6**, some recommendations are provided according to the dissertation findings. In general, our findings highlight that changing an individual habit could be difficult, but changing a whole culture is even more challenging. The research system is dynamic and involves the relationship of numerous different actors. Thus, promoting research integrity requires frequent reflection on the subject and an awareness that this discussion should be permanent.

REFERENCES

Abdi, Shila, Daniel Pizzolato, Benoit Nemery, and Kris Dierickx. 2021. "Educating PhD Students in Research Integrity in Europe." *Science and Engineering Ethics* 27 (1):5. doi: 10.1007/s11948-021-00290-0.

Afonja, Suwaiba, Damonie G. Salmon, Shadelia I. Quailey, and W. Marcus Lambert. 2021. "Postdocs' advice on pursuing a research career in academia: A qualitative analysis of free-text survey responses." *PLOS ONE* 16 (5):e0250662. doi: 10.1371/journal.pone.0250662.

Alfredo, Katherine, and Hillary Hart. 2011. "The University and the Responsible Conduct of Research: Who is Responsible for What?" *Science and Engineering Ethics* 17 (3):447-457. doi: 10.1007/s11948-010-9217-3.

ALLEA. 2017. "The European Code of Conduct for Research Integrity, revised edition." *Berlin2017 [Available from: http://www.allea.org/wp-content/uploads/2017/04/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf.*

Ampollini, Ilaria, and Massimiano Bucchi. 2020. "When Public Discourse Mirrors Academic Debate: Research Integrity in the Media." *Science and Engineering Ethics* 26 (1):451-474. doi: 10.1007/s11948-019-00103-5.

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

Armond, Anna Catharina Vieira, Bert Gordijn, Jonathan Lewis, Mohammad Hosseini, János Kristóf Bodnár, Soren Holm, and Péter Kakuk. 2021. "A scoping review of the literature featuring research ethics and research integrity cases." *BMC Medical Ethics* 22 (1):50. doi: 10.1186/s12910-021-00620-8.

Armond, Anna Catharina Vieira, and Péter Kakuk. 2021. "Research integrity guidelines and safeguards in Brazil." *Accountability in Research*:1-17. doi: 10.1080/08989621.2021.1979969.

Aubert Bonn, Noémie, and Wim Pinxten. 2019. "A Decade of Empirical Research on Research Integrity: What Have We (Not) Looked At?" *Journal of Empirical Research on Human Research Ethics* 14 (4):338-352. doi: 10.1177/1556264619858534.

Baker, Monya. 2016. "1,500 scientists lift the lid on reproducibility." *Nature* 533 (7604):452-454. doi: 10.1038/533452a.

Bauchner, H., P. B. Fontanarosa, A. Flanagin, and J. Thornton. 2018. "Scientific Misconduct and Medical Journals." *Jama* 320 (19):1985-1987. doi: 10.1001/jama.2018.14350.

Benedictus, Rinze, Frank Miedema, and Mark W. J. Ferguson. 2016. "Fewer numbers, better science." *Nature* 538 (7626):453-455. doi: 10.1038/538453a.

Bird, S. J. 2004. "Publicizing scientific misconduct and its consequences." *Sci Eng Ethics* 10 (3):435-6.

Bird, Stephanie J. 2001. "Mentors, advisors and supervisors: Their role in teaching responsible research conduct." *Science and Engineering Ethics* 7 (4):455-468. doi: 10.1007/s11948-001-0002-1.

Bouter, L. M., J. Tijdink, N. Axelsen, B. C. Martinson, and G. Ter Riet. 2016. "Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity." *Res Integr Peer Rev* 1:17. doi: 10.1186/s41073-016-0024-5.

Bülow, W., and G. Helgesson. 2019. "Criminalization of scientific misconduct." *Med Health Care Philos* 22 (2):245-252. doi: 10.1007/s11019-018-9865-7.

Cohen, Jacob. 2013. *Statistical power analysis for the behavioral sciences*: Routledge.

Collins, Caitlyn. 2020. "Productivity in a pandemic." *Science* 369 (6504):603-603. doi: doi:10.1126/science.abe1163.

Cook, Colleen, Fred Heath, and Russel L. Thompson. 2000. "A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys." *Educational and Psychological Measurement* 60 (6):821-836. doi: 10.1177/00131640021970934.

COPE Council. 2019. COPE Guidelines: Retraction Guidelines.

Copeland, P. 2022. "Stop describing academic teaching as a 'load'." *Nature*. doi: 10.1038/d41586-022-00145-z.

Crain, A. L., B. C. Martinson, and C. R. Thrush. 2013. "Relationships between the Survey of Organizational Research Climate (SORC) and self-reported research practices." *Sci Eng Ethics* 19 (3):835-50. doi: 10.1007/s11948-012-9409-0.

Curry, S. 2018. "Let's move beyond the rhetoric: it's time to change how we judge research." *Nature* 554 (7691):147. doi: 10.1038/d41586-018-01642-w.

Cyranoski, D. 2018. "China introduces 'social' punishments for scientific misconduct." *Nature* 564 (7736):312. doi: 10.1038/d41586-018-07740-z.

Davis, M. S., M. Riske-Morris, and S. R. Diaz. 2007a. "Causal factors implicated in research misconduct: evidence from ORI case files." *Sci Eng Ethics* 13 (4):395-414. doi: 10.1007/s11948-007-9045-2.

Davis, M. S., M. Riske-Morris, and S. R. Diaz. 2007b. "Causal factors implicated in research misconduct: Evidence from ORI case files." *Science and Engineering Ethics* 13 (4):395-414. doi: 10.1007/s11948-007-9045-2.

de Vries, Raymond, Melissa S. Anderson, and Brian C. Martinson. 2006. "Normal Misbehavior: Scientists Talk about the Ethics of Research." *Journal of empirical research on human research ethics : JERHRE* 1 (1):43-50. doi: 10.1525/jer.2006.1.1.43.

Dresser, R. 1997. "Giving scientists their due. The Imanishi-Kari decision." *Hastings Cent Rep* 27 (3):26-8.

DuBois, J. M., E. E. Anderson, J. Chibnall, K. Carroll, T. Gibb, C. Ogbuka, and T. Rubbelke. 2013. "Understanding research misconduct: a comparative analysis of 120 cases of professional wrongdoing." *Account Res* 20 (5-6):320-38. doi: 10.1080/08989621.2013.822248.

Else, H. 2019. "Female scientists get less money and staff for their first labs." *Nature*. doi: 10.1038/d41586-019-00933-0.

Fanelli, D. 2009. "How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data." *Plos One* 4 (5). doi: 10.1371/journal.pone.0005738.

Fanelli, D., R. Costas, F. C. Fang, A. Casadevall, and E. M. Bik. 2018. "Testing Hypotheses on Risk Factors for Scientific Misconduct via Matched-Control Analysis of Papers Containing Problematic Image Duplications." *Sci Eng Ethics*. doi: 10.1007/s11948-018-0023-7.

Fanelli, D., R. Costas, and V. Lariviere. 2015. "Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity." *PLoS One* 10 (6):e0127556. doi: 10.1371/journal.pone.0127556.

Fanelli, Daniele. 2010. "Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data." *PLOS ONE* 5 (4):e10271. doi: 10.1371/journal.pone.0010271.

Fanelli, Daniele. 2018. "Opinion: Is science really facing a reproducibility crisis, and do we need it to?" *Proceedings of the National Academy of Sciences* 115 (11):2628-2631. doi: 10.1073/pnas.1708272114.

Fang, F. C., R. G. Steen, and A. Casadevall. 2012. "Misconduct accounts for the majority of retracted scientific publications." *Proc Natl Acad Sci U S A* 109 (42):17028-33. doi: 10.1073/pnas.1212247109.

Forsberg, Ellen-Marie, Frank O. Anthun, Sharon Bailey, Giles Birchley, Henriette Bout, Carlo Casonato, Gloria González Fuster, Bert Heinrichs, Serge Horbach, Ingrid Skjæggestad Jacobsen, Jacques Janssen, Matthias Kaiser, Inge Lerouge, Barend van der Meulen, Sarah de Rijcke, Thomas Saretzki, Margit Sutrop, Marta Tazewell, Krista Varantola, Knut Jørgen Vie, Hub Zwart, and Mira Zöller. 2018. "Working with Research Integrity—Guidance for Research Performing Organisations: The Bonn PRINTEGER Statement." *Science and Engineering Ethics* 24 (4):1023-1034. doi: 10.1007/s11948-018-0034-4.

Goddiksen, M. P., U. Quinn, N. Kovács, T. B. Lund, P. Sandøe, O. Varga, and M. Willum Johansen. 2021. "Good friend or good student? An interview study of perceived conflicts between personal and academic integrity among students in three European countries." *Account Res* 28 (4):247-264. doi: 10.1080/08989621.2020.1826319.

Greenberg, D. S. 1995. "Resounding echoes of Gallo case." *Lancet* 345 (8950):639.

Grimes, David Robert, Chris T. Bauch, and John P. A. Ioannidis. 2018. "Modelling science trustworthiness under publish or perish pressure." *Royal Society Open Science* 5 (1):171511. doi: doi:10.1098/rsos.171511.

Gullifer, Judith M, and Graham A Tyson. 2014. "Who has read the policy on plagiarism? Unpacking students' understanding of plagiarism." *Studies in Higher Education* 39 (7):1202-1218.

Gunsalus, C. K. 2019. "Make reports of research misconduct public." *Nature* 570 (7759):7. doi: 10.1038/d41586-019-01728-z.

Gunsalus, C. K., Adam R. Marcus, and Ivan Oransky. 2018. "Institutional Research Misconduct Reports Need More Credibility." *JAMA* 319 (13):1315-1316. doi: 10.1001/jama.2018.0358.

Hatch, Anna, and Stephen Curry. 2020. "Changing how we evaluate research is difficult, but not impossible." *eLife* 9:e58654. doi: 10.7554/eLife.58654.

Haven, T. L., J. K. Tijdink, B. C. Martinson, and L. M. Bouter. 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):e0210599. doi: 10.1371/journal.pone.0210599.

Haven, Tamarinde L., Lex M. Bouter, Yvo M. Smulders, and Joeri K. Tijdink. 2019. "Perceived publication pressure in Amsterdam: Survey of all disciplinary fields and academic ranks." *PLOS ONE* 14 (6):e0217931. doi: 10.1371/journal.pone.0217931.

Haven, Tamarinde L., Marije Esther Evalien de Goede, Joeri K. Tijdink, and Frans Jeroen Oort. 2019. "Personally perceived publication pressure: revising the Publication Pressure Questionnaire (PPQ) by using work stress models." *Research Integrity and Peer Review* 4 (1):7. doi: 10.1186/s41073-019-0066-6.

Henriksen, Dorte. 2016. "The rise in co-authorship in the social sciences (1980–2013)." *Scientometrics* 107 (2):455-476. doi: 10.1007/s11192-016-1849-x.

Hesselmann, F., V. Graf, M. Schmidt, and M. Reinhart. 2017. "The visibility of scientific misconduct: A review of the literature on retracted journal articles." *Current Sociology* 65 (6):814-845. doi: 10.1177/0011392116663807.

Holman, L., D. Stuart-Fox, and C. E. Hauser. 2018. "The gender gap in science: How long until women are equally represented?" *PLoS Biol* 16 (4):e2004956. doi: 10.1371/journal.pbio.2004956.

Hong, S. T. 2016. "We Should Not Forget Lessons Learned from the Woo Suk Hwang's Case of Research Misconduct and Bioethics Law Violation." *J Korean Med Sci* 31 (11):1671-1672. doi: 10.3346/jkms.2016.31.11.1671.

Huang, Junming, Alexander J. Gates, Roberta Sinatra, and Albert-László Barabási. 2020. "Historical comparison of gender inequality in scientific careers across countries and disciplines." *Proceedings of the National Academy of Sciences* 117 (9):4609-4616. doi: 10.1073/pnas.1914221117.

Ioannidis, John P. A. 2005. "Why Most Published Research Findings Are False." *PLOS Medicine* 2 (8):e124. doi: 10.1371/journal.pmed.0020124.

Jensen, L. B., K. O. Kyvik, R. Leth-Larsen, and M. B. Eriksen. 2018. "Research integrity among PhD students within clinical research at the University of Southern Denmark." *Dan Med J* 65 (4).

Kakuk, P. 2009. "The legacy of the Hwang case: research misconduct in biosciences." *Sci Eng Ethics* 15 (4):545-62. doi: 10.1007/s11948-009-9121-x.

Katsarov, Johannes, Roberto Andorno, André Krom, and Mariëtte van den Hoven. 2021. "Effective Strategies for Research Integrity Training—a Meta-analysis." *Educational Psychology Review*. doi: 10.1007/s10648-021-09630-9.

Kornfeld, D. S. 2012. "Perspective: research misconduct: the search for a remedy." *Acad Med* 87 (7):877-82. doi: 10.1097/ACM.0b013e318257ee6a.

Kow, Cheryl Shumin, Yao Hao Teo, Yao Neng Teo, Keith Zi Yuan Chua, Elaine Li Ying Quah, Nur Haidah Binte Ahmad Kamal, Lorraine Hui En Tan, Clarissa Wei Shuen Cheong, Yun Ting Ong, Kuang Teck Tay, Min Chiam, Stephen Mason, and Lalit Kumar Radha Krishna. 2020. "A systematic scoping review of ethical issues in mentoring in medical schools." *BMC Medical Education* 20 (1):246. doi: 10.1186/s12909-020-02169-3.

Larivière, Vincent, Chaoqun Ni, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. 2013. "Bibliometrics: Global gender disparities in science." *Nature* 504 (7479):211-213. doi: 10.1038/504211a.

Ljubenković, A. M., A. Borovečki, M. Ćurković, B. Hofmann, and S. Holm. 2021. "Survey on the Research Misconduct and Questionable Research Practices of Medical Students, PhD Students, and Supervisors at the Zagreb School of Medicine in Croatia." *J Empir Res Hum Res Ethics* 16 (4):435-449. doi: 10.1177/15562646211033727.

Loikith, L., and R. Bauchwitz. 2016. "The Essential Need for Research Misconduct Allegation Audits." *Sci Eng Ethics* 22 (4):1027-1049. doi: 10.1007/s11948-016-9798-6.

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

Martinson, B. C., M. S. Anderson, and R. de Vries. 2005. "Scientists behaving badly." *Nature* 435 (7043):737-8. doi: 10.1038/435737a.

Martinson, B. C., A. L. Crain, R. De Vries, and M. S. Anderson. 2010a. "The importance of organizational justice in ensuring research integrity." *J Empir Res Hum Res Ethics* 5 (3):67-83. doi: 10.1525/jer.2010.5.3.67.

Martinson, B. C., L. A. Crain, R. De Vries, and M. S. Anderson. 2010b. "The importance of organizational justice in ensuring research integrity." *Journal of Empirical Research on Human Research Ethics* 5 (3):67-83. doi: 10.1525/jer.2010.5.3.67.

Martinson, B. C., C. R. Thrush, and A. Lauren Crain. 2013. "Development and validation of the Survey of Organizational Research Climate (SORC)." *Sci Eng Ethics* 19 (3):813-34. doi: 10.1007/s11948-012-9410-7.

Martinson, Brian C., David Nelson, Emily Hagel-Campbell, David Mohr, Martin P. Charns, Ann Bangerter, Carol R. Thrush, Joseph R. Ghilardi, Hanna Bloomfield, Richard Owen, and James A. Wells. 2016. "Initial Results from the Survey of Organizational Research Climates (SOuRCe) in the U.S. Department of Veterans Affairs Healthcare System." *PLOS ONE* 11 (3):e0151571. doi: 10.1371/journal.pone.0151571.

Marusic, A., E. Wager, A. Utrobicic, H. R. Rothstein, and D. Sambunjak. 2016. "Interventions to prevent misconduct and promote integrity in research and publication." *Cochrane Database Syst Rev* 4:Mr000038. doi: 10.1002/14651858.MR000038.pub2.

Marušić, Ana, Lana Bošnjak, and Ana Jerončić. 2011. "A Systematic Review of Research on the Meaning, Ethics and Practices of Authorship across Scholarly Disciplines." *PLOS ONE* 6 (9):e23477. doi: 10.1371/journal.pone.0023477.

Mejlgaard, N., L. M. Bouter, G. Gaskell, P. Kavouras, N. Allum, A. K. Bendtsen, C. A. Charitidis, N. Claesen, K. Dierickx, A. Domaradzka, A. Reyes Elizondo, N. Foeger, M. Hiney, W. Kaltenbrunner, K. Labib, A. Marušić, M. P. Sørensen, T. Ravn, R. Ščepanović, J. K. Tijdink, and G. A. Veltri. 2020. "Research integrity: nine ways to move from talk to walk." *Nature* 586 (7829):358-360. doi: 10.1038/d41586-020-02847-8.

Mills, D., and K. Inouye. 2021. "Problematizing 'predatory publishing': A systematic review of factors shaping publishing motives, decisions, and experiences." *Learned Publishing* 34 (2):89-104. doi: https://doi.org/10.1002/leap.1325.

Moher, David, Lex Bouter, Sabine Kleinert, Paul Glasziou, Mai Har Sham, Virginia Barbour, Anne-Marie Coriat, Nicole Foeger, and Ulrich Dirnagl. 2020. "The Hong Kong Principles for assessing researchers: Fostering research integrity." *PLOS Biology* 18 (7):e3000737. doi: 10.1371/journal.pbio.3000737.

Mullard, A. 2021. "Half of top cancer studies fail high-profile reproducibility effort." *Nature* 600 (7889):368-369. doi: 10.1038/d41586-021-03691-0.

Mumford, Michael D., Stephen T. Murphy, Shane Connelly, Jason H. Hill, Alison L. Antes, Ryan P. Brown, and Lynn D. Devenport. 2007. "Environmental Influences on Ethical Decision Making: Climate and Environmental Predictors of Research Integrity." *Ethics & Behavior* 17 (4):337-366. doi: 10.1080/10508420701519510.

Napoli, Philip M., and Joe Karaganis. 2010. "On making public policy with publicly available data: The case of U.S. communications policymaking." *Government Information Quarterly* 27 (4):384-391. doi: https://doi.org/10.1016/j.giq.2010.06.005.

National Academies of Sciences, Engineering, and Medicine. 2017. *Fostering integrity in research*: National Academies Press.

National Academy of Sciences, National Academy of Engineering, Responsibility Institute of Medicine Panel on Scientific, and Research the Conduct of. 1993. *Responsible Science: Ensuring the Integrity of the Research Process: Volume II*. Washington (DC): National Academies Press (US)

Copyright 1993 by the National Academy of Sciences. All rights reserved.

National Academy of Sciences, National Academy of Engineering, Responsibility Institute of Medicine Panel on Scientific, and the Conduct of Research. 1992. *Responsible Science: Ensuring the Integrity of the Research Process: Volume I*. Washington (DC): National Academies Press (US)

Copyright © 1992 by the National Academy of Sciences.

Nature. 2019. "When it comes to good practice in science, we need to think global but act local." *Nature* 576 (7786):181. doi: 10.1038/d41586-019-03782-z.

Normile, D. 2014. "Stem cell research. RIKEN panel finds misconduct in controversial paper." *Science* 344 (6179):23. doi: 10.1126/science.344.6179.23.

Nuzzo, Regina. 2014. "Scientific method: Statistical errors." *Nature* 506 (7487):150-152. doi: 10.1038/506150a.

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

OECD. 2007. "Revised field of science and technology (FoS) classification in the Frascati manual."1-12.

Opel, D. J., D. S. Diekema, and E. K. Marcuse. 2011. "Assuring research integrity in the wake of Wakefield." *Bmj* 342:d2. doi: 10.1136/bmj.d2.

Park, Chris. 2003. "In Other (People's) Words: Plagiarism by university students--literature and lessons." *Assessment & Evaluation in Higher Education* 28 (5):471-488. doi: 10.1080/02602930301677.

Paruzel-Czachura, Mariola, Lidia Baran, and Zbigniew Spendel. "Publish or be ethical? Publishing pressure and scientific misconduct in research." *Research Ethics* 0 (0):1747016120980562. doi: 10.1177/1747016120980562.

Pupovac, V., and D. Fanelli. 2015a. "Scientists Admitting to Plagiarism: A Meta-analysis of Surveys." *Sci Eng Ethics* 21 (5):1331-52. doi: 10.1007/s11948-014-9600-6.

Pupovac, V., and D. Fanelli. 2015b. "Scientists Admitting to Plagiarism: A Meta-analysis of Surveys." *Science and Engineering Ethics* 21 (5):1331-1352. doi: 10.1007/s11948-014-9600-6.

Ravn, Tine, and Mads P. Sørensen. 2021. "Exploring the Gray Area: Similarities and Differences in Questionable Research Practices (QRPs) Across Main Areas of Research." *Science and Engineering Ethics* 27 (4):40. doi: 10.1007/s11948-021-00310-z.

Redman, B. K., and J. F. Merz. 2008. "Sociology - Scientific misconduct: Do the punishments fit the crime?" *Science* 321 (5890):775. doi: 10.1126/science.1158052.

Redman, Barbara K. 2013. *Research misconduct policy in biomedicine: beyond the bad-apple approach*: MIT Press.

Resnik, D. B., and G. E. Dinse. 2013. "Scientific retractions and corrections related to misconduct findings." *J Med Ethics* 39 (1):46-50. doi: 10.1136/medethics-2012-100766.

Retraction Watch. 2015. "What should an ideal retraction notice look like?", accessed January 30. https://retractionwatch.com/2015/05/21/what-should-an-ideal-retraction-notice-look-like/.

Ščepanović, Rea, Krishma Labib, Ivan Buljan, Joeri Tijdink, and Ana Marušić. 2021. "Practices for Research Integrity Promotion in Research Performing Organisations and Research Funding Organisations: A Scoping Review." *Science and Engineering Ethics* 27 (1):4. doi: 10.1007/s11948-021-00281-1.

Schein, Edgar H. 1999. "Sense and nonsense about culture and climate."

Schwartz, Jason L. 2020. "Evaluating and Deploying Covid-19 Vaccines — The Importance of Transparency, Scientific Integrity, and Public Trust." *New England Journal of Medicine* 383 (18):1703-1705. doi: 10.1056/NEJMp2026393.

SCImago, (n.d.). "SJR — SCImago Journal & Country Rank [Portal]." accessed Feb 03. http://www.scimagojr.com.

Seglen, Per O. 1997. "Why the impact factor of journals should not be used for evaluating research." *BMJ* 314 (7079):497. doi: 10.1136/bmj.314.7079.497.

Shamsoddin, Erfan, Zahra Torkashvand-Khah, Ahmad Sofi-Mahmudi, Leila Janani, Payam Kabiri, Ehsan Shamsi-Gooshki, and Bita Mesgarpour. 2021. "Assessing research

misconduct in Iran: a perspective from Iranian medical faculty members." *BMC Medical Ethics* 22 (1):74. doi: 10.1186/s12910-021-00642-2.

Shen, H. 2013. "Inequality quantified: Mind the gender gap." *Nature* 495 (7439):22-4. doi: 10.1038/495022a.

Sijtsma, K. 2016. "Playing with Data--Or How to Discourage Questionable Research Practices and Stimulate Researchers to Do Things Right." *Psychometrika* 81 (1):1-15. doi: 10.1007/s11336-015-9446-0.

Sørensen, Mads P., Tine Ravn, Ana Marušić, Andrea Reyes Elizondo, Panagiotis Kavouras, Joeri K. Tijdink, and Anna-Kathrine Bendtsen. 2021. "Strengthening research integrity: which topic areas should organisations focus on?" *Humanities and Social Sciences Communications* 8 (1):198. doi: 10.1057/s41599-021-00874-y.

Sovacool, Benjamin K. 2008. "Exploring Scientific Misconduct: Isolated Individuals, Impure Institutions, or an Inevitable Idiom of Modern Science?" *Journal of Bioethical Inquiry* 5 (4):271. doi: 10.1007/s11673-008-9113-6.

Steneck, N. H. 2002. "Institutional and individual responsibilities for integrity in research." *Am J Bioeth* 2 (4):51-3. doi: 10.1162/152651602320957574.

Steneck, N. H. 2006. "Fostering integrity in research: definitions, current knowledge, and future directions." *Sci Eng Ethics* 12 (1):53-74.

Stern, A. M., A. Casadevall, R. G. Steen, and F. C. Fang. 2014. "Financial costs and personal consequences of research misconduct resulting in retracted publications." *Elife* 3:e02956. doi: 10.7554/eLife.02956.

Stroebe, W., T. Postmes, and R. Spears. 2012. "Scientific Misconduct and the Myth of Self-Correction in Science." *Perspect Psychol Sci* 7 (6):670-88. doi: 10.1177/1745691612460687.

Sulik, Justin, Ophelia Deroy, Guillaume Dezecache, Martha Newson, Yi Zhao, Marwa El Zein, and Bahar Tunçgenç. 2021. "Facing the pandemic with trust in science." *Humanities and Social Sciences Communications* 8 (1):301. doi: 10.1057/s41599-021-00982-9.

Thrush, C. R., J. V. Putten, C. G. Rapp, L. C. Pearson, K. S. Berry, and P. S. O'Sullivan. 2007. "Content Validation of the Organizational Climate for Research Integrity (OCRI) Survey." *J Empir Res Hum Res Ethics* 2 (4):35-52. doi: 10.1525/jer.2007.2.4.35.

Tijdink, J K, K Schipper, L M Bouter, P Maclaine Pont, J de Jonge, and Y M Smulders. 2016. "How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers." *BMJ Open* 6 (2):e008681. doi: 10.1136/bmjopen-2015-008681.

Tijdink, Joeri K., Lex M. Bouter, Coosje L. S. Veldkamp, Peter M. van de Ven, Jelte M. Wicherts, and Yvo M. Smulders. 2016. "Personality Traits Are Associated with

Research Misbehavior in Dutch Scientists: A Cross-Sectional Study." *PLOS ONE* 11 (9):e0163251. doi: 10.1371/journal.pone.0163251.

Trikalinos, Nikolaos A., Evangelos Evangelou, and John P. A. Ioannidis. 2008. "Falsified papers in high-impact journals were slow to retract and indistinguishable from nonfraudulent papers." *Journal of Clinical Epidemiology* 61 (5):464-470. doi: https://doi.org/10.1016/j.jclinepi.2007.11.019.

Viđak, Marin, Lana Barać, Ružica Tokalić, Ivan Buljan, and Ana Marušić. 2021. "Interventions for Organizational Climate and Culture in Academia: A Scoping Review." *Science and Engineering Ethics* 27 (2):24. doi: 10.1007/s11948-021-00298-6.

Wadman, Meredith. 2019. "NIH ‘high risk, high reward’ awardees skew male—again." *Science* 366 (6463):290-290. doi: doi:10.1126/science.366.6463.290.

Wakefield, A. J., S. H. Murch, A. Anthony, J. Linnell, D. M. Casson, M. Malik, M. Berelowitz, A. P. Dhillon, M. A. Thomson, P. Harvey, A. Valentine, S. E. Davies, and J. A. Walker-Smith. 1998. "Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children." *Lancet* 351 (9103):637-41. doi: 10.1016/s0140-6736(97)11096-0.

Wells, F. 2010. "The Stoke CNEP saga - did it need to take so long?" *J R Soc Med* 103 (9):352-6. doi: 10.1258/jrsm.2010.10k010.

Wells, James A., Carol R. Thrush, Brian C. Martinson, Terry A. May, Michelle Stickler, Eileen C. Callahan, and Karen L. Klomparens. 2014. "Survey of Organizational Research Climates in Three Research Intensive, Doctoral Granting Universities." *Journal of Empirical Research on Human Research Ethics* 9 (5):72-88. doi: 10.1177/1556264614552798.

Woolston, C. 2020a. "Postdoc survey reveals disenchantment with working life." *Nature* 587 (7834):505-508. doi: 10.1038/d41586-020-03191-7.

Woolston, C. 2020b. "Postdocs under pressure: 'Can I even do this any more?'." *Nature* 587 (7835):689-692. doi: 10.1038/d41586-020-03235-y.

Xie, Y., K. Wang, and Y. Kong. 2021. "Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis." *Sci Eng Ethics* 27 (4):41. doi: 10.1007/s11948-021-00314-9.

Yank, V, and D Barnes. 2003. "Consensus and contention regarding redundant publications in clinical research: cross-sectional survey of editors and authors." *Journal of Medical Ethics* 29 (2):109-114. doi: 10.1136/jme.29.2.109.

PUBLICATION LIST

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/81/2022.PL
Subject: PhD Publication List

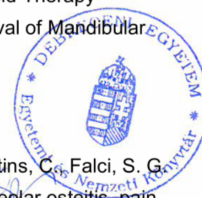
Candidate: Anna Catharina Vieira Armond
Doctoral School: Doctoral School of Health Sciences

List of publications related to the dissertation

1. **Armond, A. C. V.**, Gordijn, B., Lewis, J., Hosseini, M., Bodnár, J. K., Holm, S., Kakuk, P.: A scoping review of the literature featuring research ethics and research integrity cases.
   BMC Med Ethics. 22 (1), 1-14, 2021.
   DOI: http://dx.doi.org/10.1186/s12910-021-00620-8
   IF: 2.652 (2020)
2. **Armond, A. C. V.**, Kakuk, P.: Research integrity guidelines and safeguards in Brazil.
   Account. Res. [Epub ahead of print], 2021.
   DOI: http://dx.doi.org/10.1080/08989621.2021.1979969
   IF: 2.622 (2020)

List of other publications

3. **Armond, A. C. V.**, Glória, J. C. R., dos Santos, C. R. R., Galo, R., Falci, S. G. M.: Acupuncture on anxiety and inflammatory events following surgery of mandibular third molars: a split-mouth, randomized, triple-blind clinical trial.
   Int. J. Oral Maxillofac. Surg. 48 (2), 274-281, 2019.
   DOI: http://dx.doi.org/10.1016/j.ijom.2018.07.016
   IF: 2.068
4. Fernandes, I. A., **Armond, A. C. V.**, Falci, S. G. M.: The Effectiveness of the Cold Therapy (cryotherapy) in the Management of Inflammatory Parameters after Removal of Mandibular Third Molars: a Meta-Analysis.
   Int Arch Otorhinolaryngol. 23 (02), 221-228, 2019.
   DOI: http://dx.doi.org/10.1055/s-0039-1677755



5. **Armond, A. C. V.**, Jalles Milani, L. M., de Fátima Barbosa, F. J., de Castro Martins, C., Falci, S. G. M.: Does the use of intra-alveolar chlorhexidine gel reduces the rate of alveolar osteitis, pain, edema and trismus after the extraction of lower third molars? A meta analysis.
   J Oral Maxillofac Surg Med Pathol. 29 (6), 491-498, 2017.
   DOI: http://dx.doi.org/10.1016/j.ajoms.2017.06.012

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

6. **Armond, A. C. V.**, Martins, C. C., Glória, J. C. R., Galvão, E. L., dos Santos, C. R. R., Falci, S. G. M.: Influence of third molars in mandibular fractures. Part 1: mandibular angle-a meta-analysis.
Int. J. Oral Maxillofac. Surg. 46 (6), 716-729, 2017.
DOI: http://dx.doi.org/10.1016/j.ijom.2017.02.1264
IF: 2.164

7. **Armond, A. C. V.**, Martins, C. C., Glória, J. C. R., Galvão, E. L., dos Santos, C. R. R., Falci, S. G. M.: Influence of third molars in mandibular fractures. Part 2: mandibular condyle-a meta-analysis.
Int. J. Oral Maxillofac. Surg. 46 (6), 730-739, 2017.
DOI: http://dx.doi.org/10.1016/j.ijom.2017.02.1265
IF: 2.164

8. Glória, J. C. R., Martins, C. C., **Armond, A. C. V.**, Galvão, E. L., dos Santos, C. R. R., Falci, S. G. M.: Third Molar and Their Relationship with Caries on the Distal Surface of Second Molar: A Meta-analysis.
J. Maxillofac. Oral Surg. 17 (2), 129-141, 2017.
DOI: http://dx.doi.org/10.1007/s12663-017-1032-9

9. **Armond, A. C. V.**, Gonçalves, P. F., Flecha, O. D., de Oliveira, D. W. D., Sampaio, F. C., Falci, S. G. M.: Biosafety knowledge for the main risk activities involving public servers, students and cleaning staff of the dentistry course of UFVJM/Diamantina.
RBOL. 3 (2), 32-52, 2016.
DOI: http://dx.doi.org/10.21117/rbol.v3i2.4

Total IF of journals (all publications): 11,67
Total IF of journals (publications related to the dissertation): 5,274

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

14 February, 2022

KEYWORDS

Keywords: research integrity, research ethics, misconduct, integrity climate, publication pressure, plagiarism, questionable research practice

ACKNOWLEDGMENTS

I would like to firstly thank my supervisor Dr. Péter Kakuk, who was not only a supervisor but a mentor for me. In the past four years, he generously shared his knowledge and experiences, supporting my career advancement and celebrating my achievements, always aiming at moving me forward. I will be forever grateful. I extend my gratitude also to his family, who welcomed me and offered me nothing but kindness and support.

I would like to thank Dr. Orsolya Varga, who trusted in me and invited me to participate in her project, on which this dissertation is partially based. I am very thankful for all her support and valuable collaborations.

I would like to express my gratitude to Dr. Nóra Kovács, who became a friend and has always been supportive and helped me with the statistical analysis.

The support of Tempus Foundation and Stipendium Hungaricum is highly acknowledged.

Finally, I would like to thank my husband, family, and friends for all the support, patience, and never-ending encouragement that helped me complete my dissertation. Thank you!

APPENDICES

APPENDIX 1

Major and minor misbehavior items (Developed by Bouter LM, Tijdink J, Axelsen N, Martinson BC, ter Riet G. Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. Research integrity and peer review. 2016;1(1):17. doi:10.1186/s41073-016-0024-5).

Study design

1. Propose study questions which are clearly irrelevant [including questions that have already been or could be answered adequately by a systematic review of the literature]

2. Choose a clearly inadequate research design or using evidently unsuitable measurement instruments [which will not lead to a valid, reproducible and efficient answer to the main study question, taking into account the state-of-the-art in the field at issue]

3. Present grossly misleading information in a grant application

4. Write no or a clearly inadequate research protocol [in which essential details are lacking]

5. Ignore substantial safety risks of the study to participants, workers or environment

6. Ignore substantial risks of the expected findings for society or environment

7. Importantly change the research design during the study without disclosure [or – if applicable- without permission of sponsor, Institutional Review Board or Institutional Animal Care and Use Committee]

8. Give insufficient attention to the equipment, skills or expertise which are essential to perform the study

Data collection

9. Collect more data after noticing that the results are almost statistically significant [unless specified in a predefined adequate plan for interim analysis – also called ‘peeking’]

10. Fabricate data

11. Stop data collection earlier than planned because the results are already statistically significant [unless predefined stopping rules are implemented appropriately - also called ‘peeking’]

12. Not adhere to pertinent laws and regulations [including the laws and regulations for human and animal studies, safety regulations, good clinical practice, good laboratory practice etc.]

13. Inadequately handle or store data or (bio)materials [including archiving for an appropriate period]

14. Keep inadequate notes of the research process [with (digital) lab journals or its equivalent in other types of research]

15. Ignore basic principles of quality assurance

Reporting

16. Report on data-driven hypotheses without disclosure ['HARKing' - Hypothesizing After Results are Known - typically with a view to make results to appear more spectacular ('Chrysalis effect')]

17. Delete data before performing data analysis without disclosure

18. Selectively delete data, modify data or add fabricated data after performing initial data-analyses [in other words: falsification or fabrication of data]

19. Perform data-analyses not stated in the study protocol without disclosure

[or in predefined data-analysis plan – also called 'Significance chasing', 'P-hacking', 'data dredging', 'fishing expedition' or explorative subgroup analyses]

20. Report an incorrect downwardly rounded p-value [e.g. by reporting a p value of .054 as being less than .05]

21. Not report all study protocol-stipulated results [in the aggregate of all published reports on the study at issue]

22. Not publish a valid 'negative' study [in a form that is publicly available or accessible behind a paywall (article, report, website etc.)]

23. Report an unexpected finding as having been hypothesized from the start

24. Conceal results that contradict your earlier findings or convictions

25. Not report clearly relevant details of study methods

26. Not report replication problems

27. Selectively cite to enhance your own findings or convictions

28. Selectively cite to please editors, reviewers or colleagues

29. Selectively cite or cite your own work to improve citation metrics [e.g. Impact Factor, H-index]

30. Let your convictions influence the conclusions substantially

31. Insufficiently report study flaws and limitations

32. Spread study results over more papers than need [‘salami slicing’]

33. Duplicate publication without disclosure

34. Re-use of previously published data without disclosure [which may lead to double counting in meta-analyses]

35. Modify the results or conclusions of a study due to pressure of a sponsor [commercial or not-for-profit funder of the study]

36. Failure to disclose a sponsor of the study

37. Failure to disclose a relevant financial or intellectual conflict of interest [in publications, when reviewing grant proposals, or evaluating persons or institutions]

38. Handle existing conflicts of interest inadequately

39. Communicate results to the general public before a peer reviewed publication is available

40. Deliberately communicate findings inaccurately in the media or during presentations

41. Make no clear distinction between personal views and professional comments

Collaboration

42. Take no full responsibility for the integrity of the research project and its reports

43. Refuse to share data with bona fide colleagues

44. Turn a blind eye to putative breaches of research integrity by others

45. Refuse to respond to an allegation of a breach of research integrity

46. Use unpublished ideas or phrases of others without their permission [e.g. from reviewing manuscripts or grant applications, or from conference presentations - this is one of the forms plagiarism can take]

47. Use published ideas or phrases of others without referencing [this is one of the forms plagiarism can take]

48. Re-use parts of your own publications without referencing [‘self-plagiarism’]

49. Unfairly review papers, grant applications or colleagues applying for promotion

50. Review your own papers

51. Demand, accept or offer substantial gifts for doing a favor [e.g. authorship, promotion, access to data, favorable review or recommendation]

52. Insufficiently supervise or mentor junior coworkers

53. Be grossly unfair to your collaborators [e.g. in terms of a just balance of benefits and burdens, including giving those who deserve the opportunity to qualify as author]

54. Add an author who doesn't qualify for authorship ['honorary or gift authorship']

55. Demand or accept an authorship for which you don't qualify ['honorary or gift authorship']

56. Omit a contributor who deserves authorship ['ghost authorship']

57. Not acknowledge contributors who do not qualify for authorship

58. Not ask permission from contributors for the wording of the acknowledgement

59. Not share reviewers' comments with all co-authors

60. Submit or resubmit a paper or grant application without consent from all authors

APPENDIX 2

Origin of articles	Frequency	Percentage	Valid percentage
Missing information	81	20.51	
Australia	10	2.53	3,18
Brazil	1	0.25	0,32
Canada	15	3.80	4,78
China	5	1.27	1,59
Denmark	2	0.51	0,64
France	1	0.25	0,32
Hungary	1	0.25	0,32
India	3	0.76	0,96
Iran	2	0.51	0,64
Ireland	1	0.25	0,32
Japan	5	1.27	1,59
Korea	5	1.27	1,59
Mexico	1	0.25	0,32
Netherlands	2	0.51	0,64
New Zealand	5	1.27	1,59
Norway	2	0.51	0,64
Portugal	1	0.25	0,32
Singapore	1	0.25	0,32
South Africa	3	0.76	0,96
Spain	2	0.51	0,64
Sweden	2	0.51	0,64
Switzerland	3	0.76	0,96
Taiwan	1	0.25	0,32
Turkey	2	0.51	0,64
UK	46	11.65	14,65
USA	192	48.61	61,15
	395	100	100

APPENDIX 3

Countries of cases	Frequency	Percentage
USA	140	55.12
UK	23	9.06
Canada	14	5.51
Japan	13	5.12
China	5	1.97
Germany	5	1.97
Netherlands	4	1.57
France	3	1.18
Romania	3	1.18
Sweden	3	1.18
Australia	2	0.79
Austria	2	0.79
Croatia	2	0.79
Denmark	2	0.79
Guatemala	2	0.79
India	2	0.79
New Zealand	2	0.79
South Africa	2	0.79
Turkey	2	0.79
Uganda	2	0.79
Brazil	1	0.39
Burkina Faso	1	0.39
Congo	1	0.39
Dominican Republic	1	0.39
Ethiopia	1	0.39
Iran	1	0.39
Ivory Coast	1	0.39
Italy	1	0.39
Kenya	1	0.39
Malawi	1	0.39
Nigeria	1	0.39
Norway	1	0.39
Poland	1	0.39
Scotland	1	0.39
Singapore	1	0.39
South Korea	2	0.79
Switzerland	1	0.39
Tanzania	1	0.39
Thailand	1	0.39
Zimbabwe	1	0.39

APPENDIX 5

The study included questions on whether the participants collaborate or plan to collaborate with others as part of their research and what type of data they collect or analyze in their research. These questions help tailor the following questions according to their research context. For instance, a participant who selected quantitative data would get a statement "Not mentioning in a research publication that you replaced some outliers in a data set with data points obtained through estimates based on the remaining data points". While other who selected interviews as their type of data would get "Not mentioning in a research publication that you made up a number of direct quotes from an interview based on your general impression of the interviewee's viewpoints". Following are the full descriptions for the tailored questions:

PERCEPTIONS AND PRACTICAL JUDGMENTS ON RULES OF DATA PRACTICES

Please indicate whether you believe the following actions go against the rules and regulations that apply to you in relation to data collection, analysis and presentation.

Scenario 1:

Not mentioning in a research publication that you…

…[Quantitative data]: …removed some deviating data points from a data set when the cause of the deviation was unknown.

…[Qualitative data]: …discarded one or more interviews that did not fit well with the rest of your interviews when the cause of the deviation was unknown.

…[Artifacts/Historical]: …discarded some of your material that did not fit well with the rest of your material when the cause of the deviation was unknown.

Scenario 2:

Not mentioning in a research publication that you…

…[Quantitative data]: …removed some deviating data points from a data set when the cause of the deviation was known.

…[Qualitative data]: …discarded one or more interviews that did not fit well with the rest of your interviews when the cause of the deviation was known.

…[Artifacts/Historical]: …discarded some of your material that did not fit well with the rest of your material when the cause of the deviation was known.

Scenario 3:

Not mentioning in a research publication that you…

…[Quantitative data]: …replaced some outliers in a data set with data points obtained through estimates based on the remaining data points.

…[Qualitative data]: …made up a number of direct quotes from an interview based on your general impression of the interviewee’s viewpoints.

…[Artifacts/Historical]: …made up a number of quotes based on your general impression of the viewpoints expressed in a source.

PERCEPTIONS ON PEERS’ ATTITUDE TOWARDS RCR

To what extent do you agree with the following statement: “It is common for researchers in my field to...”

Scenario 1:

It is common for researchers in my field to...

[To all]: …allow research group leaders, supervisors or others in power to become coauthors of papers, even though they did not make a significant contribution to them.

Scenario 2:

It is common for researchers in my field to...

[Quantitative data]: …delete deviating data points based on a gut feeling that they were inaccurate.

[Non quantitative data]: …ignore parts of their material that do not fit well with the rest of the material based only on a gut feeling that the deviating material is unreliable.

Scenario 3:

It is common for researchers in my field to...

[Quantitative data]: …perform misleading or dubious statistical analysis in order to achieve publishable results.

[Non quantitative data]: …give a misleading or dubious interpretation of sources, artifacts or qualitative data in order to achieve results the teacher will accept.

Scenario 4:

It is common for researchers in my field to...

[Quantitative data]: …keep inadequate records of parts of their research that should be documented (e.g. laboratory records, descriptions of data sampling procedures, key words used in a literature study, procedures for qualitative coding, types of statistical tests performed etc.).

[Non quantitative data]: …keep inadequate records of parts of their work that should be documented (e.g. descriptions of data sampling procedures, key words used in a literature study, procedures for qualitative coding etc.).

Scenario 5:

It is common for researchers in my field to...

[To all]: …copy shorter passages from other sources into their research papers without marking them as quotes.

Scenario 6:

It is common for researchers in my field to...

[To all]: …cite sources that are not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour.

SELF-REPORTED ATTITUDE TOWARDS RCR

During your PhD, have you...

[To all]: …allowed research group leaders, supervisors or others in power to become coauthors of papers, even though they did not make a significant contribution to them.

Scenario 2:

During your PhD, have you...

[Quantitative data]: …deleted deviating data points based on a gut feeling that they were inaccurate.

[Non quantitative data]: …ignored parts of their material that do not fit well with the rest of the material based only on a gut feeling that the deviating material is unreliable.

Scenario 3:

During your PhD, have you...

[Quantitative data]: …performed misleading or dubious statistical analysis in order to achieve publishable results.

[Non quantitative data]: …given a misleading or dubious interpretation of sources, artifacts or qualitative data in order to achieve results the teacher will accept.

Scenario 4:

During your PhD, have you...

[Quantitative data]: …kept inadequate records of parts of their research that should be documented (e.g. laboratory records, descriptions of data sampling procedures, key words used in a literature study, procedures for qualitative coding, types of statistical tests performed etc.).

[Non quantitative data]: …kept inadequate records of parts of their work that should be documented (e.g. descriptions of data sampling procedures, key words used in a literature study, procedures for qualitative coding etc.).

Scenario 5:

During your PhD, have you...

[To all]: …copied shorter passages from other sources into their research papers without marking them as quotes.

Scenario 6:

During your PhD, have you...

[To all]: …cited sources that are not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour.