

SHORT THESIS 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

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

By Anna Catharina Vieira Armond, DMD, MSc

Supervisor: Dr. Péter Kakuk

Doctoral School of Health Sciences, University of Debrecen

Head of the **defence board**: Prof. Dr. Margit Balázs, Ph.D., D.Sc.

Reviewers: Dr. Béla Birkas, Ph.D.

Dr. Judit Kovács Molnárné, Ph.D.

Members of the defence board: Prof. Dr. László Fésüs, Ph.D., D.Sc., MHAS

Dr. Imre Szebik, Ph.D.

Date of defense of the dissertation: June 17, 2022, 9:00 a.m. (online format)

Publicity is provided online. If you wish to participate, please indicate it in a message sent to the e-mail address [jenei.tibor@med.unideb.hu](mailto:jenei.tibor@med.unideb.hu) by 9 a.m. on the 16<sup>th</sup> of June, the latest. After the deadline, it is no longer possible to connect to the defense due to technical reasons.

## **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. The commitment to these values by institutions and researchers represents responsible conduct of research (RCR). 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. 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. 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.

### **Concepts and definitions**

In the 1980s and 1990s, widely publicized research misconduct cases shocked the scientific community. 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. Accordingly, this narrative focused the policy efforts on the investigations and punishment of those involved in misconduct rather than preventive measures. 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". These practices have been pointed out as one of the main causes of the reproducibility crisis, 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. 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. However, with new evidence, it became clear that less serious and way more prevalent questionable research practices can be more damaging to scientific integrity.

### **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.

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. 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 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.

Furthermore, research misbehavior and QRPs might affect the reproducibility of study results, affecting research accuracy and reliability. 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.

### **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. Later this narrative moved to a broader understanding, including individual, organizational, and structural factors.

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. For instance, an empirical study by Tjebk et al. (2016) found that Machiavellianism<sup>1</sup> is associated with self-reported research misbehavior. Narrative

---

<sup>1</sup> 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.

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. 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 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. The other structural factor is the scientific award system. However, this assumption is still controversial. 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. 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) 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. Accordingly, improvements to routine research practices are critical to achieving research integrity and robustness.

Science reliability and reproducibility have been tested and questioned, and many times have found disturbing results. Although some can see the narrative of a “reproducibility crisis” as counterproductive, 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. 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 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>). 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. **Study 1** 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 **Study 2** 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 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 investigate 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.

## **MATERIALS AND METHODS**

Research integrity evidence can be investigated from different approaches and methodologies. This dissertation consists in four studies with the following methodologies:

### **Study 1 - A scoping review of the literature featuring research ethics and research integrity cases**

A 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). Articles with case(s) involving a violation of, or misbehavior, poor judgment, or detrimental research practice in relation to a normative framework, fictional and non-fictional cases 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 search was conducted in PubMed, Web of Science, SCOPUS, JSTOR, Ovid, and Science Direct in March 2018, without any language or date restrictions. 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)).

After searching the databases, 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.

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.

## **Study 2 - Characterization of PhD students and disciplinary differences on perceptions and attitude towards RCR**

A cross-sectional survey was conducted with PhD students from Hungarian universities between February and December 2020 as part of the INTEGRITY project (<https://h2020integrity.eu/>). 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). 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 anonymous online survey included demographic questions and questions on RCR perceptions, experiences, attitudes, and education and its gaps. The survey addressed clear and grey zones on three themes, citation and plagiarism, collaboration and authorship, and data practices. The survey content and the addressed topics were based on a qualitative interview study, also part of the same project.

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 squared, where  $\eta^2 = H - k + 1 / n - k$ , and the interpretation values are: 0.01- < 0.06 (small effect), 0.06 - < 0.14 (moderate effect) and  $\geq 0.14$  (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 are: 0.06- < 0.16 (small effect), 0.17 - < 0.28 (moderate effect) and  $\geq 0.29$  (large effect).

### **Study 3 - Perceptions of research integrity climate in Hungarian universities: results from a survey among academic researchers**

A cross-sectional survey study was conducted with PhD students, postdocs, and professors from three Hungarian universities (University of Debrecen, Pécs, and Szeged). The survey included demographic questions, such as gender, age, scientific field, career stage, and the Survey of Organizational Research Climate (SOuRCE).

The SOuRCE survey consists in seven subscales: 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.

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 Departmental Expectations) were calculated and stratified for academic rank and scientific field. 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 squared, where  $\eta^2 = H - k + 1 / n - k$ , and the interpretation values are: 0.01- < 0.06 (small effect), 0.06 - < 0.14 (moderate effect) and  $\geq 0.14$  (large effect). For the pairwise 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  $\geq 0.5$  (large effect). Multivariate regression analyses were performed to assess confounding interactions (with gender, age, academic rank, and scientific field) on the variables that were significant.

#### **Study 4 - Perceptions of publication pressure: survey with researchers from Hungarian universities**

A cross-sectional study was conducted with PhD students, postdocs, and professors from three Hungarian universities (University of Debrecen, Pécs, and Szeged). The survey included demographic questions, such as gender, age, scientific field, career stage, and the Revised Publication Pressure Questionnaire (PPQR). 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.

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 squared, where  $\eta^2 = H - k + 1 / n - k$ , and the interpretation values are: 0.01- < 0.06 (small effect), 0.06 - < 0.14 (moderate effect) and  $\geq 0.14$  (large effect). For the pairwise 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  $\geq 0.5$  (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

### **Study 1 - A scoping review of the literature featuring research ethics and research integrity cases**

A total of 11,641 records were identified, and 388 items were included in the qualitative synthesis. The papers contained 500 case descriptions. After applying the eligibility criteria, 238 cases were included in the analysis.

In the case analysis, fabrication and falsification were the most prevalent violations (44.9%). The lack of adherence to pertinent laws and regulations, such as lack of informed consent and REC approval, was the second most prevalent violation (15.7%), followed by patient safety issues (11.1%) and plagiarism (6.9%). 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. Paper retraction was the most prevalent sanction (45.4%), followed by exclusion from funding applications (35.5%).

The study findings show that 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 compared to its proportion in scientific publications. The cases mostly involve fabrication, falsification, and patient safety issues. This finding could have a significant impact on the academic representation of misbehaviors. The predominance of fabrication and falsification cases might diverge the attention of the academic community from relevant but less visible violations, and from recently emerging forms of misbehaviors.

## **Study 2 - Characterization of PhD students and disciplinary differences on perceptions and attitude towards RCR**

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%).

### ***Self-perception on knowledge and understanding of good practices***

The questionnaire included three questions on self-perception of good practices understandings related to the three explore themes (plagiarism and citation practices, authorship practices and collaboration, and data practices). Citation and plagiarism had the highest rates of good self-reported understanding and authorship practices had the lowest. Although there were some divergences across scientific fields, a Kruskal-Wallis test showed no significant differences

### ***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%). A Kruskal-Wallis test showed that disciplinary differences significantly affect uncertainty when facing situations involving collaboration and authorship practices ( $p=0.038$ ), 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.

### ***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”.

Related to plagiarism and citation practices, most respondents could recognize a clear misconduct scenario as a serious violation. However, this proportion decreased when introduced to QRPs or gray zone scenarios. There were no significant disciplinary differences after adjusting for confounding (integrity course attendance and gender).

In authorship practices, the scenarios concerned authorship assignment. The theme had the highest proportion of uncertainties. Only 51% of the respondents perceive it as a violation to add someone as coauthor who does not qualify to be one. Adding a supervisor who only provided critical comments on the manuscript was recognized as a violation by 36% of the respondents. Significant disciplinary differences were found between Humanities and STEMM PhD students.

In data practices, the scenarios were about data cleaning and handling outliers or discrepant data. In general, most of the scenarios were seen as violations. In the two scenarios of data cleaning, significant disciplinary differences were found between Humanities and STEMM PhD students, ( $p=0.001$ ) and ( $p=0.044$ ) for scenario 1 and 2, respectively. No differences were found in scenario 3.

### ***Perceptions on peers’ and self-reported 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. 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%).

In scenario 1, a Kruskal-Wallis test showed a statistically significant difference in the respondents’ perceptions of their peers’ attitudes regarding authorship ( $p < 0.001$ ). Post-hoc test Bonferroni-corrected shows significant differences between the Humanities and Social Sciences ( $p = 0.007$ ) and between Humanities and STEMM ( $p < 0.001$ ).

In self-reported attitude, 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”.

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%). In scenario 1, as in perceptions of peers’ attitudes, the self-reported attitude had significant differences across disciplines, according to

Fisher's exact test ( $p=0.007$ ), even after adjusted for confounding. While 32% of STEM PhD students reported having assigned authorship to someone who does not qualify to be one, only 3% of Humanities PhD students reported the same. For all the other scenarios, there were no significant disciplinary differences

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%).

### ***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.

### **Study 3 - Perceptions of research integrity climate in Hungarian universities: results from a survey among academic researchers**

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.

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 ( $\alpha$ ), ranged from 0.69 to 0.88.

### ***SOURCE© subscales and career stage***

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.

In pairwise analysis, 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,  $p=0.001$ ), and to associate or full professors (mean score= 3.43, 95% CI 3.30-3.57,  $p=0.014$ ). 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,  $p=0.027$ ), and to associate or full professors (mean score= 3.51, 95% CI 3.37-3.66,  $p=0.002$ ). 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,  $p=0.032$ ), and to associate or full professors (mean score= 3.73, 95% CI 3.61-3.86,  $p=0.006$ ). 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,  $p=0.008$ ). 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,  $p=0.007$ ). 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,  $p<0.001$ ), and with associate or full professors (mean score= 3.51, 95% CI 3.37-3.66,  $p=0.040$ ). PhD students also scored higher when compared with an associate or full professor ( $p=0.032$ ). 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.

### ***SOURCE© subscales and scientific field***

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

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,  $p=0.025$ ). 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,  $p=0.032$ ). 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 ( $p=0.046$ ). There was no evidence of a difference between the other pairs. When corrected for confounding with the career stage, the associations remained significant.

#### **Study 4 - Perceptions of publication pressure: survey with researchers from Hungarian universities**

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 ( $\alpha$ ), 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.

##### ***Publication pressure and career stage***

There was a statistically significant difference between groups in the Stress subscale as determined by the Kruskal-Wallis test. However, the adjusted pairwise analysis was not statistically significant. In Resources, both PhD students (mean score= 2.59, 95% CI 2.49-2.68,  $p<.001$ ), postdocs and assistant professors (mean score= 2.25, 95% CI 2.09-2.40,

p=.012) 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). PhD students also perceive less Resources than postdoc and assistant professors (p=.002).

### ***Publication pressure and gender***

Gender differences were found in two scales, Stress and Resources. 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), even after correction for confounding factors (age and career stage). In the Resources scale, female researchers also scored significantly higher than male researchers, but the difference was not significant after correction for confounding (career stage).

## **DISCUSSION**

### **Study 1 - A scoping review of the literature featuring research ethics and research integrity cases**

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. We collected and analyzed 507 detailed case descriptions from 392 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 mentioned cases were discussed in 132 articles.

### ***Article analysis***

Our findings show that almost all article genres discuss those cases that are new and in vogue. Misconduct and retraction notices were exceptions in the article genre analysis, as they presented mostly unique cases. However, in some cases, only a few details are provided about the case. Despite the existence of ethical guidelines, there is no uniform and common approach to how a journal should report a retraction. Moreover, 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.

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 58.8% 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.

### *Analysis of cases*

The countries from which most cases were reported were the USA (55.1%), the UK (9.0%), Canada (5.5%), Japan (5.1%), and China (1.9%). 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. However, the same interpretation does not apply to China, which occupies the second position in the ranking, with 11.2%. Our findings can 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.

The results show that 80.8% of reported cases occurred in the medical and health sciences whilst only 1.3% occurred in the humanities. 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.. 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.

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, 12.6% of cases involved issues in the study design, 16.5% in data collection, 54.5% in reporting, and 15.2% involved collaboration issues. However, as demonstrated in a meta-analysis, 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

In our study, detrimental research practices made up only 4.1% of the cases, while FFP made up 64.5%. 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.

### ***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.

## **Study 2 - Characterization of PhD students and disciplinary differences on perceptions and attitude towards RCR**

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.

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. 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. 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.

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. Studies on data practices report that researchers are often in doubt about the rules and standards applied of RCR. Data handling and cleaning are usually the most visible difficulties when interpreting the rules.

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. 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. 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%).

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 shows that 29% (95% CI 24%-35%) observed or self-reported misusing authorship.

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, 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. In the Humanities field, publications are usually single-authored and with fewer coauthors than

in other areas. Thus, unfair assignment of authorship might be less frequent, and the rules might be more transparent or less relevant.

### ***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.

### **Study 3 - Perceptions of research integrity climate in Hungarian universities: results from a survey among academic researchers**

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. and are similar to the results from Amsterdam. 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. The Hungarian findings had the worst scores on this scale, contrasting to the findings from The Netherlands and the U.S. 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.

Career stage 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 in the U.S., and contrast to the results found in Amsterdam, where PhD students scored lower than professors across all scales. The potential explanations for the low scores of postdocs and assistant professors were shorter time in the research environment and less contact to the institution. Recent surveys 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.

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 from Amsterdam, 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. 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 other similar studies, 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.

### **Study 4 - Perceptions of publication pressure: survey with researchers from Hungarian universities**

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 reliability.

Our findings show no differences in perceptions of publication pressure between scientific fields. A previous study from Amsterdam found small differences across scientific fields in Stress and Resources scales. However, the differences were greater when associated with the career stage.

In the career stage 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 career stage or scientific field. Gender differences in academia have been broadly documented from the perspectives of citations, publications, salary, recognition, and promotions. Women are more likely to get less funding, work support, and be hit harder in the work-life imbalance. Moreover, a recently published study 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.

### ***Limitations***

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. This scenario could have led to an overestimation of our results.

## **THESIS DISCUSSION**

### **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 (Tijndik, 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, perverse 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.

## NEW FINDINGS

- 1 - Scientific literature is dominated by discussions regarding prominent cases, primarily involving fabrication and falsification.
- 2 - There is an overrepresentation of biomedical research ethics and integrity cases.
- 3 - 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.
- 4 - PhD students in Hungary believe that they better understand RCR standards than they do in practice.
- 5 - Authorship assignment and data handling are the most causes of uncertainties among PhD students in Hungarian universities.
- 6 - There are disciplinary differences in understandings and attitude related to citation and data practices.
- 7 - There is a high prevalence of admitted and observed QRPs among PhD students in Hungarian universities
- 8 - The results show that postdocs and assistant professors perceived integrity climate more negatively than PhD students and full professors.
- 9 - There are disciplinary differences that should be considered when developing initiatives to foster better integrity climates.
- 10 - Responsible research conduct in Hungary is impacted by non-ideal research conditions related to resources, pressure, and interpersonal relations.
- 10- 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.
- 11- There are no disciplinary differences regarding the perceived publication pressure. However, early-career and women researchers are more exposed to publication pressure.

## 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. **Study 1** 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. **Study 2** 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. **Study 3** and **4** assessed organizational and structural factors that influence a researcher's behavior regarding RCR. **Study 3** assessed the perceptions on integrity climate and its relationship with career stage 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. **Study 4** assessed perceived publication pressure and its relationship with career stage, 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 the discussion, 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.

## PUBLICATION LIST



**UNIVERSITY of  
DEBRECEN**

**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>





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



## **ACKNOWLEDGEMENTS**

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!