



Ambiguity in Ethical Standards: Global Versus Local Science in Explaining Academic Plagiarism

Katerina S. Guba¹ · Angelika O. Tsivinskaya¹

Received: 21 May 2022 / Accepted: 13 November 2023 / Published online: 12 February 2024
© The Author(s) 2024

Abstract

The past decade has seen extensive research carried out on the systematic causes of research misconduct. Simultaneously, less attention has been paid to the variation in academic misconduct between research fields, as most empirical studies focus on one particular discipline. We propose that academic discipline is one of several systematic factors that might contribute to academic misbehavior. Drawing on a neo-institutional approach, we argue that in the developing countries, the norm of textual originality has not drawn equal support across different research fields depending on its level of internationalization. Using plagiarism detection software, we analyzed 2,405 doctoral dissertations randomly selected from all dissertations defended in Russia between 2007 and 2015. We measured the globalization of each academic discipline by calculating the share of publications indexed in the global citation database in relation to overall output. Our results showed that, with an average share of detected borrowings of over 19%, the incidence of plagiarism in Russia is remarkably higher than in Western countries. Overall, disciplines closely follow the pattern of higher globalization associated with a lower percentage of borrowed text. We also found that plagiarism is less prevalent at research-oriented institutions supporting global ethical standards. Our findings suggest that it might be misleading to measure the prevalence of academic misconduct in developing countries without paying attention to variations at the disciplinary level.

Keywords Academic plagiarism · Academic misconduct · Academic culture

✉ Katerina S. Guba
kguba@eu.spb.ru

Angelika O. Tsivinskaya
atsivinskaya@eu.spb.ru

¹ Center for Institutional Analysis of Science and Education, European University at St. Petersburg, 6/1A Gagarinskaya St., St. Petersburg, Russia 191187

Introduction

Recently, there has been renewed interest in academic misconduct as a result of switching attention from clearly forbidden forms (e.g., fabrication, falsification, and plagiarism) to questionable research practices (Biagioli et al., 2019). Researchers have documented how universities manipulate figures to improve their positions in international rankings (Espeland & Sauder, 2007), publishers launch predatory journals ready to publish virtually all submissions (Bagues et al., 2019), and individuals resort to gaming strategies like text-recycling (Horbach & Halfmann, 2019) and self-citations (Seeber et al., 2019) to enhance their chances for promotion. Research on less severe forms of gaming has produced a shift from descriptive studies conducted by social scientists concerned with such misbehavior in their communities (Biagioli et al., 2019) to the development of explanatory frameworks. As such, studies reveal the fruitfulness of employing different theoretical perspectives from organizational science (Berggren & Karabag, 2019; Hall & Martin, 2019; Walsh et al., 2019), game theory (Gall & Maniadis, 2019), and sociology (Hussinger & Pellens, 2019) to understand academic misconduct. These new theoretical accomplishments provide us with tools to refocus on studying severe forms of academic misconduct, including, particularly, plagiarism.@@@

Why do academics engage in misconduct? Although extensive research has been carried out on systematic causes (Bonn & Pinxtern, 2019), less attention has been paid to the variation in academic misconduct between research fields as most empirical studies focus on one particular discipline. Hence, relatively little is known about how the rate of plagiarism varies across different disciplines, especially in non-Western academia. We propose that discipline is only one of several systematic factors that might contribute to the presence of academic misbehavior. Because few writers have empirically presented the variation in academic misconduct between disciplines, the role of disciplines has not been considered. We suggest a neo-institutional approach to conceptualize the differences in academic plagiarism at the disciplinary level.

From a theoretical standpoint, we follow Honig and Bedi (2012), who have suggested that neo-institutional theory may support a hypothesis that predicts where and when the norm of textual originality may be less espoused by academics. Recent developments in neo-institutional theory show that global pressure to choose legitimate standards results only in a partial convergence due to strong local traditions and institutional memory (Dobbins & Knill, 2009; Hüther & Krücken, 2016; Paradeise & Thoenig, 2013; Stage, 2020). In relation to scientific ethos, we suggest that in the developing countries, the norm of textual originality has not been supported equally across different research fields. Plagiarism especially would be evident in less globalized disciplines since researchers in these disciplines might be less socialized to Western scientific norms and have experienced fewer enforcement actions compared with more globalized academic fields. Our focus on plagiarism in doctoral dissertations provides the opportunity to answer the following question: What level of plagiarism can be observed in

texts by Russian scholars that will not be peer-reviewed externally in countries where the norm of originality is prized most highly? Using Russian plagiarism detection software (Antiplagiat), we constructed representative sample of 2405 doctoral dissertations to study the role of academic disciplines in the prediction of research plagiarism. According to Russian law, higher education organizations, research institutes, and academic journals are obliged to use the software to check student papers, submitted manuscripts, and dissertations, making Antiplagiat a widely-used platform.

This study contributes to the existing literature on plagiarism that addresses the systemic factors of academic misconduct. While most research is conducted on Western countries, such as the US, UK, Australia, and Canada (Bonn & Pinxtern, 2019), we collected data from the developing country that allows us to demonstrate the relative character of scientific ethos. Specifically, we hypothesize that non-Western academics might have other views on the seriousness of academic misconduct. Previous studies of the academically less developed countries have revealed that the level of plagiarism in academia is higher when compared with Western countries (Hodges et al., 2017; Lewellyn et al., 2017; Thomas & Bruin, 2015; Vrana, 2018; Xie et al., 2021). These findings confirm Burton Clark's (1983) statement that an attachment to a global standard without borders is less evident in non-European and non-North American contexts. However, the focus on nation hides the variations that might exist at the disciplinary level as they differ in their attachment to global standards, norms, and practices. In this paper, we not only examine plagiarism across a range of different disciplines but also conceptualize the disciplinary differences in the rate of academic plagiarism depending on the level of its internationalization.

From an empirical and methodological perspective, our contribution is the representative study of academic plagiarism in Russia with the use of undirected methods in data collection. In fact, much of what has been written about plagiarism is based on direct methods when respondents are asked whether they have known colleagues who engaged in plagiarism, or editors could be asked whether they have uncovered plagiarism among submitted papers (Pupovac & Fanelli, 2015). However, the sensitivity of the topic is a critical limitation in self-reported evidence regarding misconduct (Biagioli et al., 2019). As an alternative, measuring unethical behavior by analyzing publications has the potential to overcome this limitation. The employed data collection strategy is an essential contribution because empirical research based on the direct detection of plagiarism in the entire population of scientists in a given country is rare (Bonn & Pinxtern, 2019). In this regard, we expand the previous studies of research plagiarism conducted by the activist association Dissernet (Rostovtsev, 2017); researchers analysed a large sample of Russian dissertations, however, the sample was not representing all disciplines over the entire country, and the attention was focused only on the large-scale plagiarism.

Global Versus Local Science in Explaining Plagiarism

Research on the causes of academic misconduct includes individual and systemic explanations that more recently have emphasized national and organizational variations. On the country level, the incidence of plagiarism is higher in the

developing countries (Hodges et al., 2017; Lewellyn et al., 2017; Xie et al., 2021). In his 2017 survey, Vrana (2018) demonstrated that plagiarized submissions were the second most significant problem experienced by Croatian editors. Xie et al. (2021) conducted a meta-analysis of the prevalence of academic misconduct and found that the rate of self-reported research misconduct is higher in developing countries than in the US and European countries. These empirical findings raise the following question: Why do developing countries have the highest rate of plagiarism?

The literature provides at least several possible explanations. First, there is a suggestion that there exist cultural differences regarding understanding of the notion of plagiarism (Yi et al., 2020). For example, “memorization and repetition have been very important in Chinese education and examinations, and students were encouraged to memorize and repeat literally the words of classic Confucian texts and other reputed historical persons, whereas proper citation (according to Western standards of referencing) was not emphasized” (Yi et al., 2020, p. 2). Although academics from different cultures do not diverge in understanding what constitutes obvious plagiarism (Gupta et al., 2021; Yi et al., 2022; Zhang & Zhang, 2016), there is a difference in understanding of what constitutes plagiarism in practice. Second, the prevalence of plagiarism might be related to the competition between Global North and South that pushes non-native English authors to borrow portions of other’s text to make manuscripts look more professionally written (Biagioli, 2012). While the first explanation emphasizes that different cultural experience prevents authors from fully absorbing scholarly conventions regarding proper referencing and paraphrasing, making academics engage in unintentional plagiarism (EI Bairi et al., 2022), the second explanation refers to the greater rate of tolerance to plagiarism in developing countries.

Both aspects are an issue for developing countries, given that these countries lack ethical legal documents and educational activities or such infrastructures are in the earlier stage of development. In contrast, highly developed countries have elaborated national and institutional frameworks and routinely have educated scientists about ethical behavior through workshops and courses (Pupovac et al., 2017). Even more important is that Western authors consider peer-reviewed scholarship the predominant publication standard (Honig & Bedi, 2012) that facilitates the enforcement of scientific norms. As enforcers of ethical standards, reviewers and editors are expected to provide an editorial process capable of revealing blatant forms of misconduct, including plagiarism (Horbach & Halfman, 2019). In developing countries, research fields differ in the level of absorbing Western experience given that some disciplines are local and their authors rarely engage with Western academic practices—they rarely receive educational experience in US and European universities or rarely have published in international journals, which usually provide the elaborated guidelines regarding academic misconduct.

In our main line of argument, we apply neo-institutional theory (DiMaggio & Powell, 1983; Meyer & Rowan, 1977) to the academic disciplines. The neo-institutional theory focuses on socially constructed beliefs, norms and rules that shape individual and organizational actions. The theory highlights that norms are not universal; rather they are situated in specific social and historical contexts (Shadnam &

Lawrence, 2011). This approach allows considering academic disciplines as a nested field, meaning that disciplines, as well as organizations, are embedded in the national field and others, including global fields, which are the source of legitimate models (Hüther & Krücken, 2016). Regarding the academic field, Drori and Moon (2006) cited the collapse of the Soviet bloc as an example of the reorientation of science from the dominance of communist ideology toward the Western core. At the same time, the empirical evidence suggests that contrary to the convergence thesis, global models are pragmatically transformed—resulting in dissimilar practices (Hüther & Krücken, 2016; Paradeise & Thoenig, 2013; Stage, 2020). Drori and Moon (2006) pointed out that, due to country differences in the intensity of scientific work, the level of government commitment and available resources might contribute to divergent tendencies from the global trend. Honig and Bedi (2012) studied management as a discipline developed in North America and only later spread to other countries and found a link between non-core countries and plagiarism, confirming that norms and practices may only be partially reproduced in developing countries.

The concept of nested fields lets us understand how global academic standards, including ethical ones, are reproduced in the context of developing countries. Although Western academia is a significant source of influence; however, it is not uniform for all research fields depending on the level of its internationalization. Authors from more globalized disciplines face Western standards more often as they publish internationally, making them conform to standards that reviewers enforce. However, not all disciplines are internationalized and the local embeddedness might prevent them from fully absorbing a Western standard. Some countries fully absorb the necessity to publish in international outlets (Hokka, 2018) while other communities have divided attitudes towards globalization (Gantman & Fernandez, 2016; López & Hicks, 2015). Post-Soviet sociology is the example of a discipline divided between “those who identify themselves with international and global science with those who are oriented towards predominantly local debates and audiences” (Sokolov, 2018, p. 5). Following the collapse of the Soviet Union, Western foundations supported the creation of a “global” sociology—they distributed research grants and scholarships, gave money for academic travel, and provided institutional support for Western-style private research centers and universities that espoused an opposite mission than the older public institutions (Sokolov, 2018). However, academic groups oriented toward local traditions are developed in parallel to those based upon “Western” values.

Given the relational character of academic misconduct Biagioli (2012), we propose that the necessity of following the Western ethical standard would be different for those who engage in the international community and those who prefer to stay local and avoid publishing internationally. Local groups might implement their own norms and practices as they are able to maintain their resource base. Thus, we propose our main hypothesis.

H1: *The incidence of plagiarism will be higher in less internationalized disciplines than those from more internationalized scientific fields.*

While the main focus of the paper is variance at the disciplinary level, we recognize that other structural factors might be relevant. We further suggest including organization-oriented factors to bring attention to the idea that misconduct

emerges as a result of organizational or institutional incentives (Greve et al., 2010). Following Greve et al. (2010), we propose that less plagiarism is expected at institutions that possess a concentration of the top researchers in their field. First, those institutions enjoy stronger faculty members and likely more effective oversight, and they have more to lose (Greve et al., 2010, p. 64). On the other hand, research institutions support “publish-or-perish” pressure that might stimulate authors to cut corners to satisfy increased requirements (Davis et al., 2007; Fanelli et al., 2018). The latter hypothesis was tested empirically, demonstrating that only blatant cash incentives influence the risk of scientific misconduct (Fanelli et al., 2018). Thus, we hypothesize that universities engaged in global competition should demonstrate a convergence with the model of world-class universities, including supporting global standards in their academic ethos. Based on this line of research, we propose:

H2: The incidence of plagiarism will be less evident for scholars whose defense was held at research-oriented organizations.

Data and Methods

AntiPlagiat Software

We used Russian plagiarism detection software (AntiPlagiat), a tool allowing for the comparison of selected texts with extensive text collections, particularly dissertations and academic publications available through the integration with Russian Scientific Library and Russian Scientific Electronic Library (eLIBRARY.RU). The Russian State Library contains a comprehensive digital collection of more than 919,000 full texts of Russian dissertations (Kopotev et al., 2021). eLIBRARY.RU module incorporates many categories of scientific publications, including journal articles, books, and book chapters. The main feature of eLIBRARY.RU is comprehensive coverage because the electronic library does not have strict policy for selecting scholarly periodicals (Moskaleva et al., 2018). To date, the journal catalog of eLIBRARY.RU comprises 14,551 active periodicals with 13,761 open-access sources. The total number of indexed publications is 41 million, and 15 million full texts are available at eLIBRARY.RU. Recently, AntiPlagiat was used to study academic ethics violations in Russian scientific periodicals related to the duplicated publication (Chekhovich & Khazov, 2022). Thus, AntiPlagiat is a reasonable instrument for detecting academic plagiarism.

AntiPlagiat provides a report for each document search that outlines problematic and non-problematic text matches; the latter includes identified self-citations and citations.¹ Like similar software, AntiPlagiat gives scores for each problematic match, defining the total as the share of all text overlaps found by the program except for those that the system classifies as citations and self-citations. Given that AntiPlagiat is merely text-matching software, users are advised to consult experts

¹ For more information, please review the AntiPlagiat user guide: https://docs.antiplagiat.ru/en/html/private_user_guide.html.

before deciding whether a text has been plagiarized, as the text-similarity score provides information but not a conclusion.

Sample

We focus on plagiarism in doctoral dissertations that relatively experienced scholars produce, as a doctoral dissertation in the Russian context is a higher degree that may be awarded after receiving the Candidate of Sciences. Doctor of Science is a degree that is comparable to the degree of doctor *habilitates* in some European countries. The degree requires authors to contribute significantly to a relevant discipline that should be confirmed by publishing at least ten journal articles in natural sciences and 15 articles in social sciences and humanities in addition to the dissertation manuscript. Our focus on doctoral dissertations, not on academic journal publications or books, allows us to avoid the limitation related to differences in disciplinary publication practices. An examination of journal papers, for example, would leave out other relevant publication formats; thus, dissertations help us to study plagiarism across different disciplines.

Our data were constructed in the following steps. First, we retrieved the list of all doctorate degrees (27,735) awarded in Russia between 2007 and 2015 from the official site of the national state agency responsible for overseeing academic degrees. Then we randomly selected 2600 doctoral dissertations from the pool of all dissertations defended to be searched for available text. We found 2468 texts of dissertations, which is 8.9% of all dissertations defended at that time. For further analysis, we excluded a small number of dissertations that were awarded in cultural studies and art history as their number was too low for any reliable analysis. The final sample contains 2405 dissertations written in the Russian language.² Our approach aimed to develop a representative sample of dissertations that we consider as significant contribution given that other researchers (first of all, a volunteer network called Dissernet), although checked several hundred thousand dissertations, used non-representative samples to study plagiarism in Russian dissertations (Rostovtsev, 2017).

Variables

Ideally, plagiarism should be identified by a human expert who reviews the textual similarity results and make conclusion based on the context. Human inspection is required to detect actual plagiarism; however, relying on experts might be time-consuming in the case of large samples. Most studies that estimate plagiarism based on text-matching software and additional human verification have analyzed a substantially small number of papers, fewer than 1000 (Pupovac, 2021). One possible way to deal with false positives is to use a percentage similarity cut-off as a threshold, a process broadly implemented by academic journals that receive a significant number of submissions (Manley, 2021). As such, a high value of textual similarity,

² The preliminary analysis of the sample of dissertations is available in Kopotev et al. (2021).

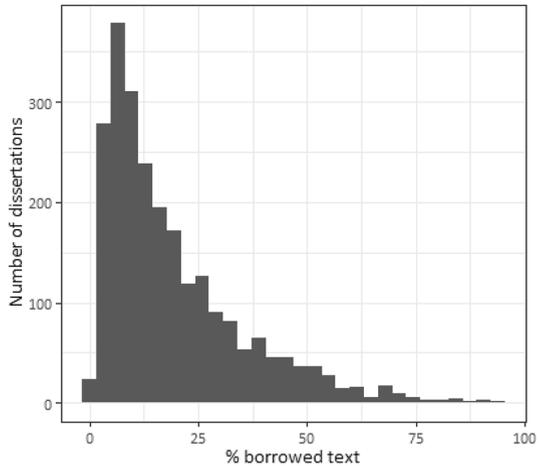
i.e., more than 50% of textual overlap, helps to eliminate false positives (Pupovac, 2021). However, there is no consensus on the acceptable extent of text similarity in scholarly articles (Gupta et al., 2021) and this method detects only the most obvious cases of text-based plagiarism. Besides, even low percentage of similarity can indicate plagiarism if a borrowed text was reproduced from one source without proper reference.

The specifics of Antiplagiat allow to consider text overlaps that do not have a proper reference as plagiarized texts given that Antiplagiat is capable, to some extent, automatically eliminate false positives, e.g., matches that are not actually plagiarism. First, it identifies text in quotes that have been cited correctly. The software defines a quotation as text in quotation marks that satisfies at least one of the conditions: (1) before the quote, the name of the author and a word denoting an utterance (e.g., speaks, writes, asserts) are present; (2) after the quotation, there is a link to the reference in square brackets or parentheses, or there is a footnote after the quote. It is critical that Antiplagiat can handle different citation styles, including cases of locating references in the footnotes allowing to consider text overlap as cases without the use of quotation marks and a precise reference to the original source. Second, Antiplagiat reviews non-problematic matches like template phrases (e.g., the names of universities and authoritative bodies, introductory words, stable speech turns, judicial acts) and matches them with collections of normative documents. Third, the software automatically identifies a bibliography (list of references), regardless of the formatting standard and language, and excludes it from analysis. Finally, Antiplagiat provides the opportunity to eliminate false positives resulting from text reuse when the software detects matches with texts published earlier by the same author. More important is the possibility of applying the special algorithm, which detects the texts published after the year in which the dissertation was defended as non-problematic.³ Thus, this algorithm signals that an author plagiarized in his/her dissertation because the cases when his/her texts were plagiarized are not taken into consideration.

In addition, Antiplagiat's setting allow an evaluator to exclude minimal matches, which might be non-problematic, by using a percentage threshold. This elimination increases the precision in labeling matched text as plagiarized text. There are different approaches to using thresholds to detect plagiarism through software, from three consecutive words to 30-word strings of text-matching (Hodges et al., 2017; Sun, 2013). For this research, a percentage of 0.01 was used to identify text we consider legitimate. On average, this percentage is equal to eight words (the mean number of words in a dissertation is 78,274). In our decision, we follow suggestions of other researchers. For example, Citron and Ginsparg (1995) stated that a text match of more than seven words might be considered problematic. Hodges et al. (2017) chose a minimum 10-word string of text-matching without attribution. Gupta et al. (2021) suggest from the personal

³ The algorithm was developed and applied by Antiplagiat's programmers specifically for this study. We are grateful to Y. Chehovich, Executive Director of Antiplagiat, whose technical assistance made it possible to obtain data about text overlap with high precision.

Fig. 1 Distribution of borrowed text



experience, a limit of at least eight to ten words should be used to avoid detecting unproblematic similarity.

Thus, we consider text overlaps that do not have a proper reference as plagiarized texts and use the percentage of borrowed text as the main dependent variable. Figure 1 presents the overall distribution of borrowed text for all dissertations in our final sample. Looking at the overall shape of the distributions (Fig. 1), we conclude that the incorrect borrowing is widespread, but cases of entire copy-paste works are rare.

For the analysis of our hypotheses, we relied on Scopus and the Russian Index of Science Citations (RISC) to collect data on the level of internationalization of academic disciplines. RISC, launched in 2005, is integrated with a full-text platform, the Scientific Electronic Library, which indexes more periodicals than RISC (Moskaleva et al., 2018). Reliance on a national source of bibliometric data is critical given that we study the non-Western higher education system, where the research is not always indexed by Web of Science or Scopus (Mongeon & Paul-Hus, 2016).

From different strategies to measure the globalization of academic disciplines, we chose to calculate the share of publications indexed in the global citation database in the overall output of academic disciplines. To calculate the number of publications produced in the global context, we queried Scopus for articles written in English with Russian affiliation in 2018 that correspond to academic disciplines in Russia (we grouped subject areas when needed and also used sub-subject areas to meet this condition). As for the number of publications produced in the local context, we relied on RISC, which covers the output of Russian journals, grouped by academic disciplines. We queried publications in Russian journals written in Russian in 2018. We limited our search to 2018, as earlier coverage of publications is known to be of less satisfying quality. Globalization was calculated as:

Table 1 Globalization by disciplines

Discipline	Globalization (%)	Median for borrowed text (%)
Physics and math	45.54	6.2
Chemistry	29.32	20.68
Biology	13.22	11.96
Earth science	11.48	8.69
Technical	9.75	14.38
Agriculture	4.79	29.12
Medicine	2.71	15.16
Economics	1.82	18.26
Philosophy	1.59	8.6
Psychology	1.11	12.71
History	1.08	10.29
Sociology	0.89	12.24
Political science	0.89	11.93
Education	0.58	15.27
Literature	0.51	6.86
Law	0.21	25.61

$$globalization_i = \frac{n_i^{Scopus}}{n_i^{Scopus} + n_i^{RICS}} \cdot 100\% \quad , i = \overline{1, m}, \quad (1)$$

where m is the number of disciplines (in our case, we considered 16), n_i^{Scopus} is the number of publications in Scopus i -th discipline, and n_i^{RICS} is the number of publications in RICS i -th discipline.

We calculated the value of globalization as highest for physics and math, with 45.5%, and lowest for the legal field, with only 0.2% (Table 1). The ordering of the academic disciplines by globalization can be considered as supporting the divide between the natural and social sciences.

In addition, the globalization of academic disciplines was calculated alternatively based only on the Russian national database, RISC, which provides access to each institution profile, including universities and research institutions. Information from the institution profile contains the overall number of publications and the number of publications indexed in the international databases Web of Science (WoS) and Scopus separately. We chose to collect information only for WoS since this citation database is known to have less representation of national journals published in languages other than English⁴ (Mongeon & Paul-Hus, 2016). Data on the number of publications was available only for six research fields: fundamental science, engineering science, medical science, social science, humanities, and agricultural science.

⁴ Around 8% out of five thousand Russian journals in RICS are indexed in WoS. These journals are considered globalized because they can be reached through international circuits of knowledge distribution.

Table 2 Globalization by fields

Field	Globalization (%)	Median for borrowed text (%)
Fundamental science	22.1	9.9
Engineering science	6.8	14.4
Medical science	6.4	15.2
Humanities	5.6	8.6
Social science	5.3	17.2
Agricultural science	3.3	29.1

As such, we were able to calculate the level of globalization only for these broad fields. Globalization was calculated as:

$$globalization_i = \frac{1}{N} \sum_{j=1}^N \frac{n_{ij}^{WoS}}{n_{ij}^{RISC}} \cdot 100\%, \quad i = \overline{1, m}, \tag{2}$$

where m is the number of fields (in our case, we considered six), n_{ij}^{WoS} is the number of publications in WoS i -th discipline by j -th organization, and n_{ij}^{RISC} is the number of publications in RISC i -th discipline by j -th organization. N is the number of organizations.

We retrieved data for about 2305 organizations registered in the RISC database with institutional profiles. The calculated value of globalization is the highest for the natural sciences field, with 22.1% (Table 2).

We used log transformation because the percentage of borrowed text has a skewed distribution with the heavy right tail. The results show that the level of globalization calculated both by disciplines and by fields is negatively correlated with the percentage of borrowed text, and it is highly statistically significant (Pearson’s correlation is $r = -0.176^{***}$, $p < 2.2e-16$ for globalization by disciplines and $r = -0.159^{***}$, $p = 4.851e-15$ for globalization by fields).

Further, we collected information on the type of institution where the defense was held. Less plagiarism is expected at institutions such as the Academy of Sciences and research universities, including participants of the Russian Excellence Program “Project 5–100.” Historically, research has been concentrated at the Academy of Sciences and not in universities. Project 5–100 was launched in 2012 to improve the international competitiveness of a handful of Russian universities. As such, universities were encouraged by financial incentives to implement the model of a world-class research university. Consequently, we coded them as research-oriented organizations and others as not research-oriented. There are 26.9% dissertations in our sample that were defended at research-oriented organizations (Table 3).

In addition, the analysis included two other independent variables as controls: gender and the defense year of a dissertation. These variables are included in line with previous research and possible variation that can be attributed to them

Table 3 Summary statistics for variables of interest

Variables	Values	Frequency (N)	Percentage (%)	% Borrowed text				
				Min	Mean	Median	Max	Std. deviation
Gender	Male	1331	55.3	0.5	19.32	14.18	91.83	16.8
	Female	1074	44.7	1.39	19.3	14.18	94.38	15.17
Defense year	2007	66	2.7	1.39	20.11	17.16	89.09	15.8
	2008	327	13.6	1.98	19.47	14.28	74.03	15.84
	2009	533	22.2	0.5	20.04	14.03	88.79	17.2
	2010	307	12.8	0.64	20.03	14.93	81.07	16.62
	2011	367	15.3	0.76	19.98	14.35	84.51	17.12
Organization type	2012	246	10.2	1.29	20.1	16.03	89.25	15.48
	2013	184	7.6	2.01	19.43	15.61	94.38	15.48
	2014	249	10.3	1.58	17.43	13.47	80.7	13.87
	2015	126	5.2	0.92	13.7	9.59	91.83	12.4
	Research-oriented	648	26.9	0.5	15.14	10.16	89.09	14.03
Discipline	Not research-oriented	1757	73.1	0.69	20.85	16.16	94.38	16.53
	Agriculture	93	3.9	3.01	31.16	29.12	84.51	18.84
	Biology	121	5.0	1.88	19.23	11.96	80.81	16.46
	Chemistry	67	2.8	1.02	25.81	20.68	81.07	19.14
	Earth science	56	2.3	0.86	11.66	8.69	57.94	11.17
	Economics	435	18.1	2.29	22.99	18.26	85.83	16.75
	History	128	5.3	1.43	14.98	10.29	91.83	14.98
	Law	106	4.4	4.53	27.86	25.61	67.61	13.51
	Medicine	329	13.7	1.98	19.18	15.16	88.79	13.94
	Education	203	8.4	2.54	20.27	15.27	89.25	15.95
Literature	116	4.8	0.84	9.18	6.86	61.04	8.52	

Table 3 (continued)

Variables	Values	Frequency (N)	Percentage (%)	% Borrowed text				
				Min	Mean	Median	Max	Std. deviation
	Philosophy	91	3.8	12.18	8.6	47.86	10.06	
	Physics and math	164	6.8	12.55	6.2	94.38	15.33	
	Political science	40	1.7	16.06	11.93	45.55	11.02	
	Psychology	52	2.2	17.53	12.71	62.89	14.3	
	Sociology	46	1.9	14.38	12.24	41.51	9.6	
	Technical	358	14.9	19.81	14.38	74.18	17.18	
	Total	2405	100					

Fig. 2 Distribution of borrowed text by gender

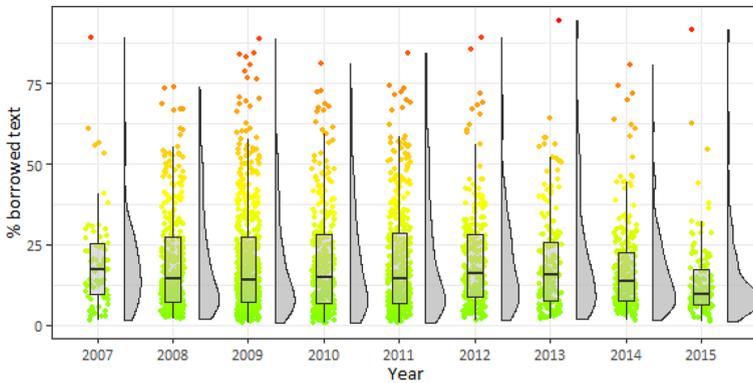
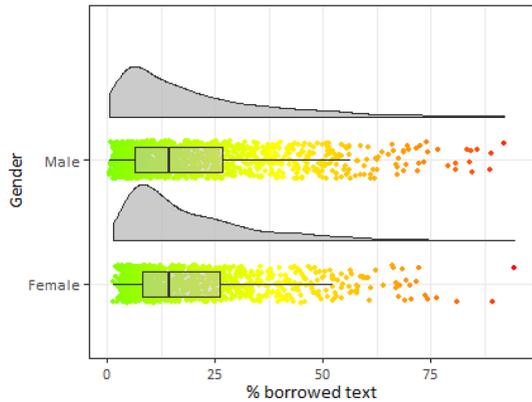


Fig. 3 Distribution of borrowed text by year

in the models. Gender is considered to be a possible relevant factor since males might be more inclined to take risks than females, and might therefore be more likely to engage in academic misconduct (Fanelli et al. (2018) proposed the gender hypothesis although the results of the study did not support it). Table 3 presents the gender distribution of doctorate holders in our sample. There are more male authors (55.3%) than female ones. We also included the defense year of a dissertation as, over the years, some reformation of the system has occurred during a ten-year span in regards to the process of obtaining the doctoral degree, making it more difficult to falsify works and raising the quality of defended theses. Defense year varies between 2007 and 2015. This period was chosen due to the availability and sufficient coverage of information on dissertations.

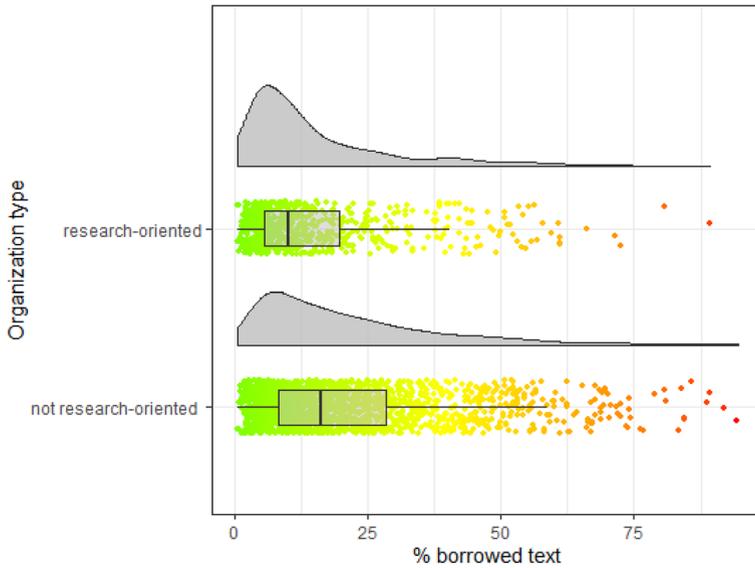


Fig. 4 Distribution of borrowed text by organization type

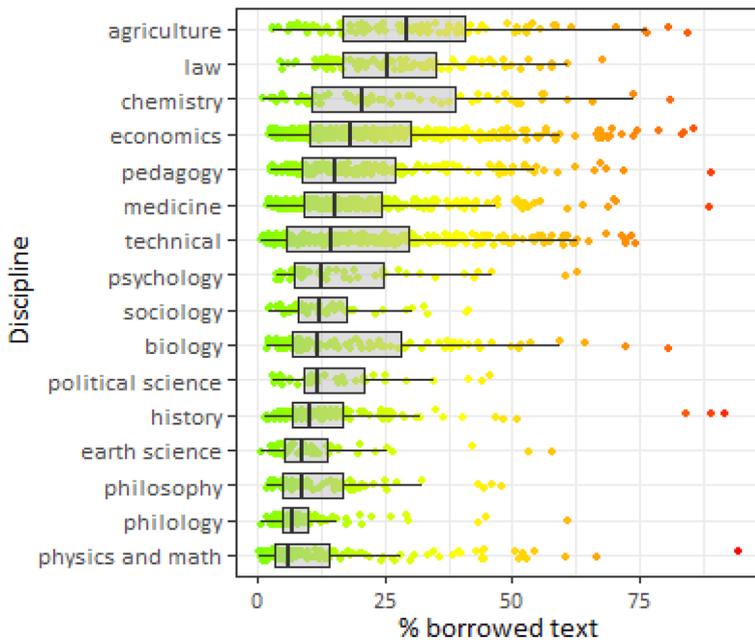


Fig. 5 Distribution of borrowed text by discipline, ordered by median percentage of borrowed text

Results

Descriptive Statistics

As a starting point for our analysis, we turn to the graphical approach to show its key features and better depict the nature of our variables. Using this data visualization approach, we show differences in the percentage of borrowed text for selected variables such as gender, the year of defense, the organization type, and discipline (Figs. 2, 3, 4, and 5).

Our data demonstrate that the average share of detected borrowings was 19.1% (median: 13.9%). Only a quarter of dissertations have a level of plagiarized text less than 7%. At the same time, it is rather unusual to copy the entire text: in the sample of 2468 dissertations, we found only 41 where borrowing exceeded 50%.⁵ Figure 2 illustrates that the distributions of borrowed text for males and females are similar in shape. There are slightly more observations, with more than 75% of borrowed text, for men. Both men and women have a median value of 14.18% with the standard deviation is 16.8% for men and 15.17% for women.

We considered the defense year of the dissertation as another common control variable (Fig. 3). The biggest difference in the percentage of borrowed text is for 2007 compared to other years (the median is 9.59%). The Mann–Kendall Trend Test for median percentage of borrowed text showed ($\tau = -0.333$, 2-sided p value = 0.25) that a trend is not statistically significant.

Regarding the influence of organizational type on the percentage of the detected plagiarized text, we can distinguish between research-oriented and not research-oriented organizations (Fig. 4). The difference in median values is around 6% less borrowed text at research-oriented organizations. The distribution for research-oriented organizations is less spread out and has a higher peak. There is some evidence for the effect on borrowed text by organization type where the dissertation was defended.

Finally, we show how academic discipline affects the distribution of borrowed text (Fig. 5). The highest median percentage of borrowed text is in agricultural dissertations (29.12%), legal sciences (25.6%), and chemical sciences (20.8%). The lowest is in physics and math (6.2%), literature (6.9%), and philosophy (8.4%) that is line with the observation made on the study of text-recycling in four research areas—Horbach and Hoffman (2019) found that humanities are disciplines for which the wording is the essence of the novelty. We found that in the social sciences, the incidence of plagiarism is evident for economics, education, psychology, and political science.

⁵ The results raise the question of how the considerable number of dissertations with the significant amount of plagiarized text were defended. By formal documents, the organizations with dissertation councils have been required to use antiplagiarism software since 2012, and until this moment, a dissertation might not have been checked at all. Even when the practice of obligatory checking was introduced, the councils might not have evaluated the results with proper attention; instead, they might have used the threshold (10 or 20%) to detect whether a dissertation contains plagiarized text. Using thresholds is widely spread among Russian academics despite the Antiplagiat's instruction that this practice might lead to erroneous conclusions.

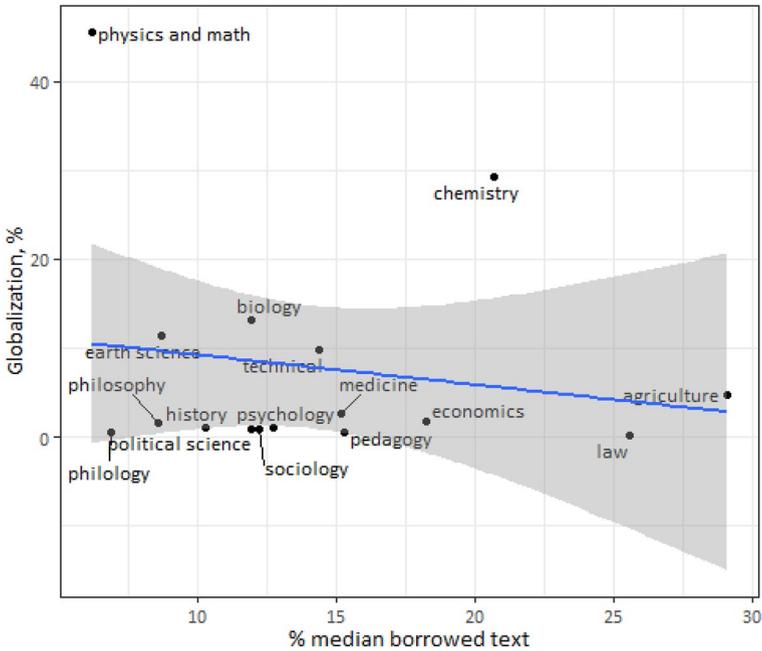


Fig. 6 Median % of borrowed text by globalization of discipline

Our results generally do not contradict the previous attempts to study plagiarism in Russian dissertations. Thus, Dissernet, an activist association, found large-scale plagiarism in over 6500 dissertations and conclude that plagiarized texts are primarily produced in economics, pedagogy, and law, followed by medical sciences, political sciences, engineering, and social sciences. In contrast, plagiarized dissertations are rare in natural sciences (Rostovtsev, 2017).

As our primary variable of interest is globalization, we engaged in a closer investigation of the relationship between globalization and the percentage of borrowed text (Fig. 6). Physics and math stand out as having the highest level of globalization and the lowest median value for borrowed text. Dissertations in chemistry have a higher median percentage compared to other natural disciplines, which are more globalized than the social sciences. Other disciplines that more closely follow the pattern of higher globalization are associated with a lower percentage of borrowed text.

Regression Models

Values of our dependent variable can be expressed as bounded by the interval (0,1). We considered several regression models, which are generally the appropriate choice for proportional data. Our selection suggested that the CDF-quantile model best describes the data, as it out-performed the other ones such as beta and fractional models. This

model has additional properties compared to alternatives as it not only distinguishes the effect median difference but also the effect of independent variables on dependent variables in terms of the variability of values, as it has two submodels (location and dispersion). There is a relation between pairs of these distributions. These pairs are “quantile-duals” of one another in the sense that one’s CDF is the other’s quantile, with the appropriate parameterization. Separate submodels may be specified for the location and dispersion parameters, with different or overlapping sets of predictors in each (Smithson et al., 2017).

We present the general equation for the CDF-quantile regression model for random variables on the (0, 1) interval (Smithson et al., 2017). Consider a random sample of N independent observations from (Y, \mathbf{V}) , where Y is a random variable and \mathbf{V} is a vector of predictors. A model for the distribution of Y , conditional on \mathbf{V} , has two submodels: the “location submodel” and the “dispersion submodel.” The location submodel has a vector \mathbf{W} of predictors selected from \mathbf{V} , and the dispersion model has another vector of predictors \mathbf{Z} of predictors selected from \mathbf{V} . The sets of predictors in \mathbf{W} and \mathbf{Z} may not overlap. For the i -th observation, these submodels may be written as:

$$\begin{aligned} L_\mu(\hat{\mu}_i) &= \mathbf{w}_i^T \boldsymbol{\beta} \\ L_\sigma(\hat{\sigma}_i) &= \mathbf{z}_i^T \boldsymbol{\delta} \end{aligned}, \quad i = \overline{1, N}. \tag{3}$$

For submodels in the CDF-quantile regression, different families of fitted distributions may be chosen. There are no clear criteria for the selection process, but based on the log-likelihood measure, our choice fell on T2-T2 distributions.

The full model includes control variables and both the globalization and the organization type in the location submodel and constant representing the spread of distribution for the response variable in the dispersion submodel (specified in Eq. 4):

$$\begin{aligned} \hat{\mu}_i &= \beta_0 + \beta_1 gender_i + \beta_2 defense_year_i + \beta_3 organization_type_i + \beta_4 globalization_i \\ \log(\hat{\sigma}_i) &= \delta_0 + \delta_1 globalization_i. \end{aligned} \tag{4}$$

The results of regression analysis with the specified model in Eq. 4 for two calculations of globalization by fields and by disciplines are presented in Table 4. In the descriptive analysis, we presented variables that might influence the percentage of borrowed text separately and in the aggregate form. In contrast, models below contain data on the individual level, taking into account the influence of variables together.

Regarding the controls, there is no statistically significant gender difference in the percentage of plagiarized text in the model by fields and a low statistically significant difference in the models by disciplines. Regarding the defense year, the trend test was not statistically significant; the only statistically significant difference in the models is between 2015 and 2007. The median value at the beginning of the period was 17.16%. In the next two years, it lowered to around 14%, and in 2012, it was higher with a median of 16.03%, while at the endpoint of the observed period, it was 9.59%. The beginning and endpoint of the period have a smaller number of dissertations, so we cannot reach a definitive conclusion about this difference and a slowing trend in the percentage plagiarized over the years. Overall, no statistically significant observable trend or pick differences are associated with possible interventions in a particular year.

Table 4 CDF-quantile regression models

	By fields	By disciplines
Location submodels		
Intercept	- 1.556 (0.138)***	- 1.662 (0.133)***
<i>Control variables</i>		
2008	- 0.111 (0.135)	- 0.106 (0.137)
2009	- 0.118 (0.13)	- 0.109 (0.131)
2010	- 0.046 (0.136)	- 0.021 (0.138)
2011	- 0.128 (0.135)	- 0.111 (0.136)
2012	- 0.029 (0.137)	- 0.007 (0.138)
2013	- 0.106 (0.142)	- 0.091 (0.143)
2014	- 0.207 (0.136)	- 0.192 (0.138)
2015	- 0.546 (0.152)***	- 0.529 (0.153)***
male	0.064 (0.045)	0.078 (0.046)*
<i>Variables of interest</i>		
Globalization	- 1.744 (0.45)***	- 1.039 (0.25)***
Not research-oriented	0.458 (0.056)***	0.467 (0.055)***
Dispersion submodels		
Intercept	- 0.209 (0.022)***	- 0.208 (0.022)***
Num. obs	2405	2405
Log likelihood	1649.491	1650.465

The results show that with the increase in the level of globalization calculated by disciplines, the percentage of borrowed text lowers. The model shows a negative coefficient in the location submodel (estimated coefficient = - 1.039, $p < 0.001$, which is the median difference. In case of globalization calculated by fields, the higher is the level of globalization, the lower is the percentage of borrowed text (estimated coefficient = - 1.744, $p < 0.001$). In both calculations, the coefficients for globalization in the models are highly statically significant. Thus, our hypothesis H1 about globalization is confirmed based on the regression analysis results. We also found a highly statistically significant difference between research-oriented and not research-oriented organizations, with the median values of 16.16% and 10.16%, respectively. Models show that the median value is lower by 0.467 (0.458) for research-oriented organizations with globalization calculated by disciplines (fields), indicating that the H2 hypothesis is also confirmed. Dissertations associated with research universities and the Russian Academy of Science have, on average, fewer plagiarized texts (compared to median values).

Discussion and Conclusion

This study was designed to determine the role of academic disciplines in addition to individual and organizational factors relevant for explaining the rate of plagiarism. To study academic plagiarism, we constructed unbiased sample of Russian doctoral dissertations which were analyzed through an online plagiarism detection service. The results suggest that the rate of incorrect borrowing is substantial among Russian

degree holders. With the threshold at 20% plagiarized text, more than a third of dissertations (37%) were flagged that is remarkably higher than in Western countries [at this level of threshold, the study showed that 5% of publications from the United States and United Kingdom were problematic (Citron & Ginsparg, 1995)]. This result is in line with recent studies which also reveal that the prevalence academic misconduct in developing countries is significantly higher than in highly developed countries (Pupovac et al., 2017; Xie et al., 2021). The empirical results showed remarkable variation across disciplines: while the incidence of incorrect borrowing is relatively low in physics, earth science, and humanities, scientists from social sciences, chemistry, and applied fields have engaged in plagiarism more regularly.

We proposed that socialization with a scientific ethos would be different for those who engaged in the international community in comparison to those who preferred to stay local. We measured the globalization of academic disciplines by calculating the share of publications indexed in the global citation database in the overall output of academic disciplines. In general, disciplines closely follow the pattern: a higher level of globalization is associated with a lower percentage of borrowed text (with chemistry as an exemption, which expresses an unexpected pattern of a globalized field with a high level of borrowed text). Even after running regression models with a number of controls, the effect is significant. We also found that that text borrowing is less prevalent in research-oriented institutions supporting global standards in their academic ethos.

We suggest that further research on structural factors explaining academic misconduct requires special efforts to improve the level of conceptualization of disciplinary variance. Horbach and Haifmann (2019) demonstrated that the extent of text-recycling is associated with differences in publication cultures among scientific disciplines. Seeber et al. (2019) considered the role of disciplinary differences in social structure to explain the incidence of self-citations as an opportunistic response to the use of metrics. However, less attention has been paid to the specifics of disciplinary conventions regarding dominant academic norms. Authors with plagiarized dissertations might not perceive themselves as scientists involved in the blatant disregard of norms as they might not prioritize the norm of complete originality or even have alternative norms. They could rationalize their actions, for example, by borrowing technical parts of the text (descriptions of experiments) or using the work of colleagues to which they have contributed (as when a supervisor writing a doctoral dissertation uses dissertations of candidates of sciences defended under their supervision).

Regarding the empirical prospect, it might be fruitful to examine whether academic disciplines express different patterns of incorrect text borrowing. Researchers usually use the percentage of borrowed text as the primary dependent variable without any attempts to categorize different types of plagiarism. We suggest that online software provides the opportunity to deepen our understanding of academic plagiarism if we focus on the practice of text borrowing. The relevant research questions include the following: Which parts of text are borrowed more often? Where are they located? How long are they? Do scientists copy substantial sections, or rather segments of small size? We suggest that the remarkable difference between disciplines might be evident regarding the different types of academic plagiarism. It is

also possible that dominant rationalizations that scientists use to explain why their behaviors differ from socially accepted patterns are reflected in the dominant practice of text borrowing.

The prevalence of academic misconduct in developing countries is still an underdeveloped topic (Vrana, 2018; Zhang, 2010); although their output in global science is increasing, national policies on research integrity are still being created. Plagiarism in non-Western countries is usually considered in the context of authors' attempts to publish something in international journals (Biagioli, 2012). Struggling with this challenge, they borrow different portions of articles to get through the standard of an Anglophone paper. However, we studied dissertations written in a local language published before the pressure to publish internationally. It seems that plagiarism in developing countries is related to more than the desire to cut corners and publish internationally at any cost. The general attitude to the norm of originality might be different, especially in disciplines (social sciences and applied fields) that do not follow the global publication culture. Local conventions to the originality of academic texts exist outside of opposition to the global norm. Studying academic misconduct in developing countries only through English-language publications might distract researchers' attention from explanations why scientists may ignore the global norm of originality.

International publishing requires absorbing strict ethical standards regarding scholarly conventions of proper referencing. Authors and editors may be less familiar with traditional scholarly conventions in developing countries. Research shows that although global editors have expressed a mainstream view on ethical standards, the attitude of non-Anglophones is a little less rigorous than that of the Anglophones (Zhang & Zhang, 2016). The international experience that varies at the level of academic disciplines allows one to get familiar with standards and reproduce more mainstream views on academic research behavior. Even if we studied one national case, the same pattern might exist in other contexts, e.g., authors engaged in international publishing are more exposed to other ethical standards that result in changing local attitudes. However, the exact answer requires comparative data from other developing countries that might be a fruitful topic for further research.

The findings may be limited by relying on the similarity of checking software which has several inevitable drawbacks. First, the plagiarism rate might be underestimated due to false negatives—the cases of plagiarism not detected by the Antiplagiat software. The obvious source for false negatives is translational plagiarism, when an academic text published in English or other languages is translated into the local language and then used with minor changes without reference to the original source. Unfortunately, these cases are extremely difficult to detect, although different approaches to address them have been recently discussed (Memon, 2020). In this regard, our results might be biased towards revealing higher plagiarism rates for dissertations from less internationalized scientific disciplines in which authors rely primarily on Russian sources. In contrast, authors from more globalized disciplines might use translational plagiarism more often, but the software was not designed to detect it. Besides, authors might use various instruments to avoid the detection of similar texts, including synonym replacement or substituting letters from the Russian alphabet with similar letters from other alphabets. However, Antiplagiat

specialists regularly monitor and counteract these techniques to inflate a paper's originality percentage.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bagues, M., Sylos-Labini, M., & Zinovyeva, N. (2019). A walk on the wild side: 'Predatory' journals and information asymmetries in scientific evaluations. *Research Policy*, *48*, 462–477. <https://doi.org/10.1016/j.respol.2018.04.013>
- Berggren, C., & Karabag, S. F. (2019). Scientific misconduct at an elite medical institute: The role of competing institutional logics and fragmented control. *Research Policy*, *48*, 428–443. <https://doi.org/10.1016/j.respol.2018.03.020>
- Biagioli, M. (2012). Recycling texts or stealing time? Plagiarism, authorship, and credit in science. *International Journal of Cultural Property*, *19*, 453–476. <https://doi.org/10.1017/S0940739112000276>
- Biagioli, M., Kenney, M., Martin, B. R., & Walsh, J. P. (2019). Academic misconduct, misrepresentation and gaming: A reassessment. *Research Policy*, *48*, 401–413. <https://doi.org/10.1016/j.respol.2018.10.025>
- Bonn, A., & Pinxten, W. (2019). A decade of empirical research on research integrity: What have we (not) looked at? *Journal of Empirical Research on Human Research Ethics*, *14*, 338–352. <https://doi.org/10.1177/1556264619858534>
- Citron, D. T., & Ginsparg, P. (2015). Patterns of text reuse in a scientific corpus. *Proceedings of the National Academy of Sciences*, *112*, 25–30. <https://doi.org/10.1073/pnas.1415135111>
- Chekhovich, Y. V., & Khazov, A. V. (2022). Analysis of duplicated publications in Russian journals. *Journal of Informetrics*, *16*(1), 101246. <https://doi.org/10.1016/j.joi.2021.101246>
- Clark, B. R. (1983). *Higher education system: Academic organisation in cross-national perspective*. California U.P.
- Davis, M. S. (2003). The role of culture in research misconduct. *Accountability in Research*, *10*, 189–201. <https://doi.org/10.1080/714906092>
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*. <https://doi.org/10.2307/2095101>
- Dobbins, M., & Knill, C. (2009). Higher education policies in Central and Eastern Europe: Convergence toward a common model? *Governance*, *22*, 397–430. <https://doi.org/10.1111/j.1468-0491.2009.01445.x>
- Drori, G. S., & Moon, H. (2006). The changing nature of tertiary education: Neo-Institutional perspectives on cross-national trends in disciplinary enrolment 1965–1995. In D. P. Baker & A. W. Wiseman (Eds.), *The impact of comparative education research on institutional theory* (pp. 163–192). Emerald. [https://doi.org/10.1016/s1479-3679\(06\)07008-3](https://doi.org/10.1016/s1479-3679(06)07008-3)
- El Bairi, K., El Kadmiri, N., & Fourtassi, M. (2022). Exploring scientific misconduct in Morocco based on an analysis of plagiarism perception in a cohort of 1,220 researchers and students. *Accountability in Research*. <https://doi.org/10.1080/08989621.2022.2110866>
- Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, *113*, 1–40. <https://doi.org/10.1086/517897>

- Fanelli, D., Costas, R., Fang, F. C., Casadevall, A., & Bik, E. M. (2018). Testing hypotheses on risk factors for scientific misconduct via matched-control analysis of papers containing problematic image duplications. *Science and Engineering Ethics*, 25, 771–789. <https://doi.org/10.1007/s11948-018-0023-7>
- Gall, T., & Maniadiis, Z. (2019). Evaluating solutions to the problem of false positives. *Research Policy*, 48, 506–515. <https://doi.org/10.1016/j.respol.2017.12.005>
- Gantman, E. R., & Fernández Rodríguez, C. J. (2016). Literature segmentation in management and organization studies: The case of Spanish-speaking countries (2000–10). *Research Evaluation*, 5, rvv031. <https://doi.org/10.1093/reseval/rvv031>
- Greve, H. R., Palmer, D., & Pozner, J.-E. (2010). Organizations gone wild: The causes, processes, and consequences of organizational misconduct. *The Academy of Management Annals*, 4, 53–107. <https://doi.org/10.1080/19416521003654186>
- Gupta, L., Tariq, J., Yessirkepov, M., Zimba, O., Misra, D. P., Agarwal, V., & Gasparyan, A. Y. (2021). Plagiarism in non-anglophone countries: a cross-sectional survey of researchers and journal editors. *Journal of Korean Medical Science*, 36(39), 21. <https://doi.org/10.3346/jkms.2021.36.e247>
- Hall, J., & Martin, B. R. (2019). Towards a taxonomy of research misconduct: The case of business school research. *Research Policy*, 48, 414–427. <https://doi.org/10.1016/j.respol.2018.03.006>
- Hodges, A., Bickham, T., Schmidt, E., & Seawright, L. (2017). Challenging the profiles of a plagiarist: A study of abstracts submitted to an international interdisciplinary conference. *International Journal for Educational Integrity*, 13, 102. <https://doi.org/10.1007/s40979-017-0016-3>
- Hokka, J. (2018). What counts as ‘good sociology’? Conflicting discourses on legitimate sociology in Finland and Sweden. *Acta Sociologica*, 62(4), 357–371. <https://doi.org/10.1177/0001699318813422>
- Honig, B., & Bedi, A. (2012). The fox in the hen house: A critical examination of plagiarism among members of the academy of management. *Academy of Management Learning & Education*, 11, 101–123. <https://doi.org/10.5465/amle.2010.0084>
- Horbach, S. P. J. M., & Halfman, W. (2019). The extent and causes of academic text recycling or ‘self-plagiarism.’ *Research Policy*, 48, 492–502. <https://doi.org/10.1016/j.respol.2017.09.004>
- Hussinger, K., & Pellens, M. (2019). Guilt by association: How scientific misconduct harms prior collaborators. *Research Policy*, 48, 516–530. <https://doi.org/10.1016/j.respol.2018.01.012>
- Hüther, O., & Krücken, G. (2016). Nested organizational fields: Isomorphism and differentiation among European universities. *Research in the Sociology of Organizations*, 46, 53–83. <https://doi.org/10.1108/s0733-558x20160000046003>
- Kopotev, M., Rostovtsev, A., & Sokolov, M. (2021). Shifting the norm: the case of academic plagiarism detection. *The Palgrave Handbook of Digital Russia Studies*. https://doi.org/10.1007/978-3-030-42855-6_27
- Lewellyn, K. B., Judge, W. Q., & Smith, A. (2017). Exploring the questionable academic practice of conference paper double dipping. *Academy of Management Learning and Education*, 16, 217–236. <https://doi.org/10.5465/amle.2015.0033>
- Lopez Pineiro, C., & Hicks, D. (2015). Reception of Spanish sociology by domestic and foreign audiences differs and has consequences for evaluation. *Research Evaluation*, 24, 78–89. <https://doi.org/10.1093/reseval/rvu030>
- Manley, S. (2021). The use of text-matching software’s similarity scores. *Accountability in Research*, 13, 1–27. <https://doi.org/10.1080/08989621.2021.1986018>
- Memon, A. R. (2020). Similarity and plagiarism in scholarly journal submissions: bringing clarity to the concept for authors, reviewers and editors. *Journal of Korean Medical Science*. <https://doi.org/10.3346/jkms.2020.35.e217>
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83, 340–363. <https://doi.org/10.1086/226550>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106, 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- Moskaleva, O., Pisylyakov, V., Sterligov, I., Akoev, M., & Shabanova, S. (2018). Russian index of science citation: Overview and review. *Scientometrics*, 116, 449–462. <https://doi.org/10.1007/s11192-018-2758-y>
- Paradeise, C., & Thoening, J.-C. (2013). Academic institutions in search of quality: Local orders and global standards. *Organization Studies*, 34, 189–218. <https://doi.org/10.1177/0170840612473550>

- Pupovac, V. (2021). The Frequency of plagiarism identified by text-matching software in scientific articles: A systematic review and meta-analysis. *Scientometrics*, 126(11), 8981–9003. <https://doi.org/10.1007/s11192-021-04140-5>
- Pupovac, V., & Fanelli, D. (2015). Scientists admitting to plagiarism: A meta-analysis of surveys. *Science and Engineering Ethics*, 21, 1331–1352. <https://doi.org/10.1007/s11948-014-9600-6>
- Pupovac, V., Prijčić-Samaržija, S., & Petrovečki, M. (2017). Research misconduct in the Croatian scientific community: A survey assessing the forms and characteristics of research misconduct. *Science and Engineering Ethics*, 23, 165–181. <https://doi.org/10.1007/s11948-016-9767-0>
- Rostovtsev, A. (2017). Plagiarism in the dissertations and scientific publications in Russia. In *Plagiarism across Europe and beyond 2017. Conference proceedings* (pp. 107–112).
- Seeber, M., Cattaneo, M., Meoli, M., & Malighetti, P. (2019). Self-citations as strategic response to the use of metrics for career decisions. *Research Policy*, 48, 478–491. <https://doi.org/10.1016/j.respol.2017.12.004>
- Shadnam, M., & Lawrence, T. B. (2011). Understanding widespread misconduct in organizations: An institutional theory of moral collapse. *Business Ethics Quarterly*, 21(3), 379–407. <https://doi.org/10.5840/beq201121324>
- Smithson, M., & Shou, Y. (2017). CDF-quantile distributions for modelling random variables on the unit interval. *British Journal of Mathematical and Statistical Psychology*, 70, 412–438. <https://doi.org/10.1111/bmsp.12091>
- Sokolov, M. (2018). The sources of academic localism and globalism in Russian sociology: The choice of professional ideologies and occupational niches among social scientists. *Current Sociology*, 67, 818–837. <https://doi.org/10.1177/0011392118811392>
- Stage, A. K. (2020). Are national university systems becoming more alike? Long-term developments in staff composition across five countries. *Policy Reviews in Higher Education*, 4, 68–104. <https://doi.org/10.1080/23322969.2019.1702088>
- Sun, Y.-C. (2013). Do journal authors plagiarize? Using plagiarism detection software to uncover matching text across disciplines. *Journal of English for Academic Purposes*, 12, 264–272. <https://doi.org/10.1016/j.jeap.2013.07.002>
- Thomas, A., & de Bruin, G. P. (2015). Plagiarism in South African management journals. *South African Journal of Science*, 111, 1–3. <https://doi.org/10.17159/sajs.2015/20140017>
- Vrana, R. (2018). Editorial challenges in a small scientific community: Study of Croatian editors. *Learned Publishing*, 31, 369–374. <https://doi.org/10.1002/leap.1188>
- Walsh, J., Lee, Y.-N., & Tang, L. (2019). Pathogenic organization in science: Division of labor and retractions. *Research Policy*, 48, 444–461. <https://doi.org/10.1016/j.respol.2018.09.004>
- Xie, Y., Wang, K., & Kong, Y. (2021). Prevalence of research misconduct and questionable research practices: A systematic review and meta-analysis. *Science and Engineering Ethics*, 27(4), 41. <https://doi.org/10.1007/s11948-021-00314-9>
- Yi, N., Nemery, B., & Dierickx, K. (2020). Perceptions of plagiarism by biomedical researchers: An online survey in Europe and China. *BMC Medical Ethics*, 21(1), 1–16. <https://doi.org/10.1186/s12910-020-00473-7>
- Yi, N., Nemery, B., & Dierickx, K. (2022). Do biomedical researchers differ in their perceptions of plagiarism across Europe? Findings from an online survey among leading universities. *BMC Medical Ethics*, 23(78). <https://doi.org/10.1186/s12910-022-00818-4>
- Zhang, H. Y. (2010). CrossCheck: An effective tool for detecting plagiarism. *Learned Publishing*, 23, 9–14. <https://doi.org/10.1087/20100103>
- Zhang, Y., & Zhang, Y. (2016). Differences between Anglophone and non-Anglophone journals. In *Against plagiarism: A guide for editors and authors* (pp. 13–38).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.