publications

Opinion

Exploring the Need to Use "Plagiarism" Detection Software Rationally

Petar Milovanovic [1], Tatjana Pekmezovic [2] and Marija Djuric [1,*]

[1] Center of Bone Biology, Institute of Anatomy, Faculty of Medicine, University of Belgrade, 11000 Belgrade, Serbia; drpmilovanovic@gmail.com
[2] Institute of Epidemiology, Faculty of Medicine, University of Belgrade, 11000 Belgrade, Serbia; pekmezovic@sezampro.rs
* Correspondence: marijadjuric5@gmail.com

Abstract: Universities and journals increasingly rely on software tools for detecting textual overlap of a scientific text with the previously published literature to detect potential plagiarism. Although software outputs need to be carefully reviewed by competent humans to verify the existence of plagiarism, university and journal staff, for various reasons, often erroneously interpret the degree of plagiarism based on the percentage of textual overlap shown in the similarity report. This is often accompanied by explicit recommendations to the author(s) to paraphrase the text to achieve an "acceptable" percentage of overlap. Here, based on the available literature and real-world examples from similarity reports, we provide a classification with extensive examples of phrases that falsely inflate the similarity index and argue the futility and dangers of rephrasing such statements just for the sake of reducing the similarity index. The examples provided in this paper call for a more reasonable assessment of text similarity. To fully endorse the principles of academic integrity and prevent loss of clarity of the scientific literature, we believe it is important to shift from pure bureaucratic and quantificational view on the originality of scientific texts to human-centered qualitative assessment of the manuscripts, including the software outputs.

Keywords: text similarity; plagiarism detection; "tortured" phrases; paraphrasing; assessment

check for updates

Academic Editor: Andrew Kirby

Received: 4 November 2024
Revised: 5 December 2024
Accepted: 30 December 2024
Published: 2 January 2025

Citation: Milovanovic, P., Pekmezovic, T., & Djuric, M. (2025). Exploring the Need to Use "Plagiarism" Detection Software Rationally. *Publications*, *13*(1), 1. https://doi.org/10.3390/publications13010001

1. Introduction

With the growing digitalization, unprecedented availability of online literature, and increasing awareness of plagiarism in submitted texts and theses at various levels of the academic and research arena, universities and journals increasingly rely on software tools for detecting textual overlap of a scientific text with the previously published literature to identify potential plagiarism. Despite growing interest in this topic, there are still concerns in practice about the advantages and disadvantages of such tools in detecting plagiarized content.

In this paper, we discuss some common misconceptions and misinterpretations of the outputs of software tools for detecting textual similarity; highlight the situations where a high percentage of textual overlap with the previous literature is predominantly a consequence of the legitimate use of common terms and phrases in scientific research and particular research field, rather than a sign of plagiarism; emphasize various strategies that are often used by students and scholars to evade software detection and falsely reduce textual overlap, such as paraphrasing, back translation, and word spinning, both in the context of a plagiarism cover-up and in an attempt to achieve linguistic diversification or address poor language abilities; discuss the place of artificial intelligence (AI) in this field, both as a deception tool and a tool to detect plagiarism when standard plagiarism detection software fails to do so; discuss dangers and futility of paraphrasing common scientific terms or phrases at all cost just to reduce the textual similarity; and finally, offer several suggestions regarding the interpretation and use of software outputs and improving the assessment of plagiarism. Of note, this paper is primarily concerned with the scientific disciplines where particular, standardized language is required to describe particular phenomena, such as in biomedicine, natural sciences, and engineering.

2. Use of Software Tools for Detecting Textual Similarity: Avoiding Misconceptions and Misinterpretations

The introduction of various software tools for detecting textual overlap of a scientific text with the previously published literature has greatly influenced the scientific community and publishing industry. Much of the value of such tools is in preventing direct copying of the work of others, and during the last two decades, scholars have been increasingly educated to pay close attention to such issues. The leading plagiarism screening services that are currently used are Turnitin and iThenticate—both products from the same company—which are currently used by numerous teachers to evaluate student classwork (Turnitin) and by universities and publishers to evaluate textual similarity in scholarly contributions (iThenticate). These tools are based on the comparison of a submitted text with numerous documents available on the Internet and provide outputs that indicate the percentage of textual overlap between the examined document and the previously published material available to the software tool. The percentage of overlap is also known as the similarity index.

In recent years, universities and journals have increasingly relied heavily on such software tools, often with an awareness of the limitations of such an approach (Radikė & Camm, 2022). However, although the companies behind such software solutions are careful in describing the intentions and capabilities of the software and prefer using the term "similarity" or "overlap" instead of "plagiarism"—leaving final interpretation to human assessment—still, many scholars and some university/journal staff colloquially use the term "extent of plagiarism" (Higgins et al., 2016) instead of "percentage of similarity" and erroneously interpret the degree of originality/plagiarism from the percentage shown in the similarity report (Baskaran et al., 2019; Manley, 2023; Pourret, 2024). In practice, this means that a low similarity index is often interpreted as a sign of originality and lack of plagiarism, whereas a high similarity index is typically interpreted as plagiarism. Moreover, the numerical value of the similarity index is usually perceived as being in direct correlation with the extent or degree of plagiarism (Manley, 2023; Pourret, 2024). Given that it is impossible to avoid any overlap with the previous literature, for practical reasons, some journals define a certain numerical threshold, suggesting that there is a certain "acceptable" level of textual overlap. The threshold may differ among journals, publishers, and universities (Mahian et al., 2017), being anywhere in the range between 10% and 30%. However, when the similarity index of a submitted manuscript exceeds this arbitrary threshold, authors are typically asked to revise the text through paraphrasing so as to reduce the similarity index.

We want to emphasize the wrongness of taking the similarity index for granted by describing illustrative examples contributing to a falsely high similarity index and some examples of "tortured phrases" that appear in the literature in an attempt to deceive software and reduce textual overlap through paraphrasing.

3. When a Number Deceives: Classification of Common Causes of a Falsely Inflated Similarity Index

Our analysis of a number of similarity reports for medical journal submissions and submitted PhD theses has shown examples where most of the overlap originated from the overlap with title page information, mandatory statements (e.g., ethics, disclosures, funding), and authors' or committee members' names and affiliations. Although it is meaningless to include such parts in the evaluation, this sometimes happens even in highly ranked journals, contributing to falsely higher similarity indexes. It is reasonable to assume that this happens by error, often in an attempt to submit a manuscript or thesis for software-based evaluation in the shortest time possible.

In addition, a higher similarity index is often mainly a consequence of the overlap with various common phrases or standard terms used in scientific communications (Table 1). These include general statements used to describe research; common phrases used in the field of medicine and common anatomical terms and phrases; description of study subjects; names of surveys or organizations; names of machines and software with the required information about their manufacturer and technical details; standard terms related to clinical or high-resolution research imaging; standard names and abbreviations for the measured parameters; names of cellular molecules, elements, or biochemical processes; common phrases related to description of statistical analysis; and standard phrases used during discussion (Table 1). While textual overlap is often less heavily judged when detected in the introduction, materials and methods, and discussion sections, it is usually considered fully unacceptable in the results section. However, even in the results section, there are some additional common phrases—such as the phrases related to reporting of statistical analysis, common phrases used in the figure legends, accidental overlap of *p* values or other coefficients, and standard phrases related to groups and intergroup comparisons—all of which may contribute to a falsely high similarity index (Table 1).

Table 1. Classification of common phrases or standard terms used in scientific communications that often falsely inflate the similarity index with examples from real-life situations.

Types of Common Phrases or Standard Terms	Examples
Statements used to describe research	*"aim was to investigate effects"; "…was divided into four phases"; "in the second part of the study"; "objectives of the present study"; "has been extensively studied"; "the overall aim of this thesis was to"*
Phrases used in the field of medicine	*"development of preventive strategies"; "high blood glucose levels"; "sports-related injuries are"; "clinical and research conditions"*
Anatomical terms and phrases	*"anterior cruciate ligament and posterior cruciate ligament"; "articular surfaces of the femur and tibia"; "open and closed kinematic chain"*
Description of study subjects	*"individuals with type 2 diabetes mellitus"; "women above the age of"; "week-old C57BL/6 J male mice were"*
Names of surveys or organizations	*"National Health and Nutrition Examination Survey (NHANES)"; "according to Surveillance data from the National Collegiate Athletic Association"*
Names of machines and software with the required information about their manufacturer and technical details	*"LEO435 VP; LEO Electron Microscopy Ltd., Cambridge, UK"; "using inductively coupled plasma–mass spectrometry [ICP-MS, iCAP Qc, Thermo Scientific, UK]"*

Table 1. *Cont.*

Types of Common Phrases or Standard Terms	Examples
Terms related to clinical or high-resolution research imaging	*"scanning parameters were as follows"; "ring artifact and beam hardening corrections"; "region of interest (ROI)"/"volume of interest (VOI)"*
Names and abbreviations for the measured parameters	*"number of entries in closed arms"; "advanced glycation end products (AGEs)"; "bone mineral density (BMD) of the lumbar spine"*
Names of cellular molecules, elements, or biochemical processes	*"calcium/calmodulin dependent protein kinase II"; "depolarization evoked neurotransmitter release"; "readily releasable pool of vesicles"*
Phrases related to description or reporting of statistical analysis	*"Student's t-test or Mann–Whitney U test, depending on. . . "; "using Pearson's or Spearman's correlation tests"; "evaluate inter-rater reliability"; "all statistical analyses were"; "using SPSS software. . . "; "with a significance level set at 0.05"; "interaction between time and . . . was not significant"; "did not show significant effect of. . . "; "was independently associated with"; "associated with an increased risk of"; "a positive correlation with"*
Phrases used in the figure legends	*"Error bars represent the standard error of the mean"*
Accidental overlap of *p* values or other coefficients	*"$r = 1$, $p > 0.05$"*
Phrases related to groups and intergroup comparisons	*"the groups were comparable"; "in comparison to the control group"*
Phrases used during discussion	*"performed a cross-sectional study"; "investigated the impact of. . . on"; "is reasonable to speculate that"; "to the best of our knowledge, there are no studies that"; "were unable to establish a causal relationship"*

Particularly illustrative examples were two theses, one in which much of the similarity stemmed from using the sentence "*Error bars represent the standard error of the mean*" in almost 40 figure legends and another with extensive statistical analysis that presented the phrase "*interaction between time and . . . was not significant*" more than 50 times, which falsely inflated the similarity index.

4. Strategies Used to Evade Software Detection and Falsely Reduce Textual Overlap (Similarity Index): Paraphrasing, Back Translation, and Word Spinning

While the so-called plagiarism detection tools have certainly helped to screen for and identify many instances of plagiarism, have stimulated scientists to acknowledge the work of others through correct citation practice, and likely had a preventive role in avoiding verbatim copying of someone else's text (Mostofa et al., 2021), they have also "stimulated" the development of various strategies to evade plagiarism detection, such as paraphrasing, back translation, and word spinning (Akbari, 2021; Alvi et al., 2021; Jones & Sheridan, 2015).

Currently, a number of paraphrasing tools are often used by students to modify someone else's text and present it as their own when they need to submit required essays or manuscripts for formal purposes (Ansorge et al., 2021). Some researchers also use these tools for various reasons. The most severe issue is the *use of such tools with the intention to deceive*, namely, to plagiarize someone else's scientific text or part of it and present it as their own work. This is, of course, a huge and important problem in the academic world and publishing industry, and there is no dilemma about its wrongness. Ansorge et al. detected an article that was basically a paraphrased version of a previously published scientific paper on the topic of "water literacy" (Ansorge et al., 2021). To prove that the paraphrased version was produced by using a paraphrasing tool, Ansorge et al. used the suspected paraphrasing tool on the original article and generated a paraphrased text very similar to the incriminated paper (Ansorge et al., 2021).

In addition to paraphrasing, translating to another language and then again back to English (back translation or back-and-forth translation (Jones, 2009; Jones & Sheridan, 2015)) by translator software tools could yield similar results (Ansorge et al., 2021); that is, it could produce a modified version of a paper or a certain paragraph, deceiving both the plagiarism detection software and readers that the text is not plagiarized. However, Prentice and Kinden (Prentice & Kinden, 2018) compared a medical text modified by the use of six free online paraphrasing tools and six separate translations using the Google Translate™ tool and showed that the paraphrasing tools often altered standard medical terminology, whereas the translation tool mostly retained standard medical terminology and nomenclature. In other words, if standard medical terms are often replaced by unidiomatic terms or phrases, it is more likely that the plagiarist used paraphrasing tools than translation-based plagiarism strategies. Nevertheless, as recently pointed out by Ansorge, translation software may also sometimes create unidiomatic terms in certain fields of science (e.g., in water research), but the appearance of those nonstandard terms should not automatically be judged as unethical (Ansorge, 2024). Namely, many non-English authors may write their papers in their mother language and then use translation software to have them translated into English. If the authors are not familiar with English terminology, do not sufficiently know the subject matter, or do not carefully check and rectify the results of AI-based translation, it is also possible to generate terms that deviate from standard English terminology used in a certain discipline, without any intention to deceive (Ansorge, 2024). To examine whether translation can also generate "tortuous phrases", Aronson used Google Translate to translate the phrase "artificial intelligence" into 121 different languages and then translate it back into English and found that in 10 of the 121 languages, "tortured phrases" appeared (e.g., "creative wisdom", "clever fraud detection", "fraudulent search") (Aronson, 2021).

In an attempt to deceive journals and universities, some authors also use so-called "word spinning software" tools, managing to fully evade detection by plagiarism detection tools. Such interventions may dramatically degrade the quality of the original text, which may even lead to fully incomprehensive and unidiomatic statements. Specifically, Kannangara conducted a study at the Waiariki Institute of Technology and examined various assignments submitted by students (Kannangara, 2017). The authors found notable degradation of text quality by word spinning, along with total evasion of plagiarism detection (Kannangara, 2017).

5. AI: Both Part of a Problem and Part of a Solution

In principle, standard plagiarism detection tools provide a similarity report with quantitative information about textual similarity or overlap and show parts of the text that match the previously published text, along with the possible sources (Memon, 2020). However, authors increasingly use various paraphrasing strategies, especially substitution by synonyms, reordering of words in sentences, and deletion and/or insertion of words or phrases (Alvi et al., 2021), thereby managing to avoid detection by current plagiarism detection tools (Kannangara, 2017). In addition, the expansion of the field of AI creates new concerns as to whether the text has been developed or modified by authors (humans) or machines.

Another potential concern is what happens to the manuscripts uploaded to various plagiarism detection software tools. Specifically, there are some concerns as to whether the uploaded manuscript would become a part of a large database, leading to marking the same manuscript as plagiarized when uploaded by a next journal following rejection in a previous one. Moreover, there are dilemmas regarding whether feeding manuscripts to software companies may create a possibility for the work to be stolen or plagiarized by AI, especially if feeding the manuscript to freely available, problematic software tools with poor ethical standards and lack of transparency. Indeed, there has been some concern over the ingestion of authors' work by the generative AI models, as some publishers have even made deals with AI companies granting them access to the published content (Potter, 2024). Nevertheless, further evidence is needed to substantiate or disprove these concerns.

Given the growing use of generative AI to write texts, there have been attempts to detect machine-generated or machine-paraphrased text, and recent studies have coined the term "tortured phrases" to account for "unspecific jargon or confusing alternative phrases" (Cabanac et al., 2021; Teixeira da Silva, 2023) created by AI writing tools.

In many cases, "tortured phrases" are caused by poor translation (Teixeira da Silva, 2023), insufficient English proficiency, or deliberate paraphrasing, whereas in other cases, especially when overly abundant in a scientific text, these phrases may even point to serious ethical issues (e.g., use of a paper mill, fabrication of paper, generation of a pseudo-scientific manuscript by online tools) (Cabanac et al., 2021; Else, 2021; Lay et al., 2022; Teixeira da Silva, 2023). Unfortunately, standard plagiarism detection software services are helpless in these cases (Kannangara, 2017), and there is growing research interest in developing other approaches based on AI to identify cases of machine-generated or machine-paraphrased text (Alvi et al., 2021; Gangadharan et al., 2020; Vrbanec & Meštrović, 2020, 2023; Wahle et al., 2022). Some of the promising strategies may include "context matching and pre-trained word embeddings" to detect the most common types of paraphrasing, namely substitution with synonyms and reordering of words (Alvi et al., 2021), and "pre-trained word embedding models combined with machine learning classifiers and state-of-the-art neural language models" (Wahle et al., 2022).

Recently, there have been new developments to detect so-called "tortured acronyms" and suspicious phrases in scientific papers—such as "Problematic Paper Screener" (https://dbrech.irit.fr/pls/apex/f?p=9999:24, accessed on 3 December 2024), which can be used by publishers to signal misconduct (O'Grady, 2024). Based on such initiatives, thousands of published papers and conference proceedings have been flagged (O'Grady, 2024). While this will likely help publishers, and maybe also universities, to screen for suspicious texts, it is also expectable that AI tools for writing will evolve to avoid such obvious signs of deception, namely the so-called "tortured phrases" and acronyms (e.g., "*convolutional brain organization*" instead of "*convolutional neural network*"; "*glucose bigotry*" instead of "*glucose intolerance*" (O'Grady, 2024); "*flag to clamor*" instead of "*signal to noise*" (Martel et al., 2024); "*well designed system that has bad performance is of no use to any man, woman or animal*" (Cabanac & Labbé, 2021)). Some tools are advertised to be able to analyze a text and give a detailed report on any sections that may be AI-generated or rephrased (e.g., QuillBot's AI Detector) or even create or modify a text with allegedly no clues of AI-generated content and that "passes Turnitin" (e.g., AI Text Humanizer or StealthGPT). While universities and publishers are trying to handle plagiarism issues, AI tools have arrived to further complicate the question of originality and plagiarism or deception.

6. Dangers and Futility of Paraphrasing Common Scientific Terms or Phrases at All Cost to Reduce the Similarity Index

Of course, not every instance of a textual overlap and not every use of paraphrasing or word-spinning tools are necessarily linked to intentional deception and plagiarism. So-called "unintentional plagiarism" is far more common (Martínez-López et al., 2019) and may be related to ignorance or unawareness of the citation rules in academic writing, as well as insufficient English proficiency. Various text intervention tools are sometimes used by authors (typically non-native English speakers) to diversify their linguistic abilities without a clear intention to deceive anyone. In this article, we particularly focus on the instances where software reports indicate a falsely inflated similarity index, which originates from the use of common phrases or terms. In such cases, when where there is actually no plagiarism, authors often try to paraphrase some common and established scientific phrases or descriptions, such as those listed in Section 3 and Table 1. This may be done out of fear that each textual overlap will be regarded as plagiarism, but it is often even required by university staff or journal editors, where authors may be advised to "change the text through paraphrasing to achieve an acceptable level of similarity/overlap" or "revise any full or nearly full sentences of highlighted text in order to reduce as much of the direct overlap as possible."

Regardless of the reason, the use of paraphrasing or word spinning tools often creates instances of poor language quality, including meaningless or nonexistent words or phrases (e.g., "*momentum water use*" instead of "*current water use*"; "*compelled to oversee in trouble*" instead of "*forced to manage in difficulty*"; "*singular contrasts*" instead of "*individual differences*"; "*advanced period*" instead of "*modern era*"—some examples extracted from the paraphrased text reported in (Ansorge et al., 2021)). Indeed, artificial interventions with an attempt to reduce the similarity index often introduce very bizarre and confusing terms not in standard use (e.g., "*overwhelming metals*" or "*substantial metals*" instead of "*heavy metals*" (Pourret, 2024)), especially when it comes to standard medical terms (e.g., "*feline output*" instead of "*CAT scan*"; "*aldohexose levels*" instead of "*glucose levels*"; "*crisis center*" or "*crisis office*" instead of "*emergency department*"; "*sort 1*" or "*kind 1*" instead of "*type 1* (diabetes)"; "*restoration focus*" instead of "*rehabilitation center*"; "*release report*" instead of "*discharge summary*"; "*healing facility*" instead of "*hospital*" (Prentice & Kinden, 2018)), which distorts the original meaning, reduces clarity, and definitely hampers the integrity of the scientific literature (Pourret, 2024).

We believe it makes no sense to reword common phrases listed in Section 3 and Table 1 just to reduce the similarity index. With the growing amount of scientific literature and a growing databases available as source to such software tools for such software tools, it will be even less possible to avoid overlap of this kind, and the use of paraphrasing tools induced by the necessity to reduce the similarity index by any means may seriously degrade the quality, clarity, and understandability of the scientific text.

7. Recommendations for Universities and Publishers That Evaluate the Originality of Scholarly Work

While we certainly do not support direct copying of someone else's text, and we fully endorse the key principles of academic integrity, we advocate for reasonable assessment of text similarity. It should be noted that situations of falsely inflated similarity index, when not interpreted correctly by the journal or university staff, may cause serious anxiety among authors, especially those who are not fully informed about the details of text analysis by the software. This category frequently includes PhD candidates who try to publish articles that are the basis for the defense of their doctoral theses (Milovanović et al., 2023). In the last instance, the author of a text with a falsely inflated degree of textual overlap may be unjustifiably exposed to ethical review, which is a source of stress and may affect career development (Milovanović et al., 2023).

In this context, we summarize some recommended attitudes and approaches for the key players in the publication process:

Universities are an important pillar in this matter, as they have both the possibility and the duty to educate students and future scientists. The following recommendations could be made for universities:

- It is important to educate students on the key principles of academic integrity and familiarize them with the anti-plagiarism policy of the university (Morán & Carlos, 2022);
- It has been suggested to assign a special department to record cases of detected and sanctioned plagiarism and implement a software tool for screening the students' works (Morán & Carlos, 2022);
- It is important to introduce education and training programs on the topics of intellectual property and academic integrity among the students and raise awareness on this subject (Devlin, 2006; Morán & Carlos, 2022; Perkins et al., 2020). This may be often performed in collaboration with librarians (Gunnarsson et al., 2014);
- We believe that judgment of originality should never rely solely on software outputs and quantitative indicators but should always include competent human reviewers to ensure correct interpretation of the originality of a scientific text, especially when this is important for career development (e.g., submission of PhD thesis for evaluation). In this context, one also has to keep in mind that the percentage of overlap depends on the software used (Mahian et al., 2017) and the language of the submitted text (Milovanović et al., 2023), and definitely, there should be no decisions based solely on automatic detection, but they should be based on additional human assessment (Bretag & Mahmud, 2009; Memon, 2020).

Publishers and journals are key stakeholders in scientific publishing and are also those that most often deal with issues of plagiarism and copyright. The following recommendations could be made for the publishing industry:

- We acknowledge that publishers (or any other large institutions) that manage thousands of submissions often set an arbitrary threshold similarity index to judge the presence of plagiarism. Such software tools are practical screening tools to flag potential plagiarism risks. However, the outputs from these tools showing textual overlap should not automatically judge the manuscript as containing plagiarized content. The decision should be made after expert analysis of the similarity report, keeping in mind the legitimate use of common phrases and standard terms such as those listed in Table 1;
- We assume that editors may face challenging situations where the publisher has a strict threshold policy for textual overlap, and the editorial assessment of the similarity report shows that the similarity index is falsely high, mostly based on the legitimate use of general phrases or standard terms and descriptions. In such situations, the editors may be obliged to request the authors to reduce the similarity index through paraphrasing or other manipulations so as to satisfy the publisher's policy. We believe that publishers should have their policies updated to account for such situations so as to facilitate workflow for editors in such situations without the need to ask the authors to paraphrase just to meet the arbitrary threshold of originality.
- The abundance of nonstandard terms in a scientific text may sometimes be a sign of fabricated research or plagiarized content. As not all cases of "tortured" phrases are signs of deliberate fabrication and intention to deceive, such cases should be flagged, but the final verdict regarding the origin and nature of such content should not be rushed. As suggested by Ansorge (Ansorge, 2024), manuscript reviewers should look

for instances of “incorrect terminology” (such as “tortured phrases”), and if they find such phrases, they should explicitly point it out in the comments to the editor and ask the authors to explain why the incorrect terminology appears in the manuscript;

- AI-based tools to detect fabricated, plagiarized, machine-written, paraphrased, or translated content should be further developed. For example, AI-based tools could be used to detect instances of “tortured” phrases in a given text, but the results should always be analyzed seriously by a competent human expert;
- It has been suggested to assess the entire publishing process from value and ethical perspectives (McCuen, 2018), where honesty and fairness should be highly endorsed, and it should be kept in mind that we require a “collective effort and commitment from authors, reviewers, editors, and policy-makers … to address the problem of plagiarism, especially in the developing and non-English speaking countries” (Memon, 2020).

8. Final Remarks

In this article, we call for a more nuanced, human-centered process of evaluating the originality of scholarly contributions that balances academic integrity with practical realities. Indeed, whenever humans leave the decision on the presence of plagiarism or lack thereof to artificial, quantitative thresholds in software-generated similarity reports, focusing on the “acceptable” similarity index (Mahian et al., 2017), we fear that we are missing the point and we are encouraging “easy fixes” reflected in the use of paraphrasing or word spinning tools, thereby eventually contributing to the loss of literature integrity, clarity, and quality.

We should never forget the essential tasks of a scientific text, namely, to provide original and reliable information in a clear, precise, and sufficiently detailed fashion to allow for understanding and repeating the experiments/analyses. This also includes the use of standard, widely accepted terminology, nomenclature, and phrases—especially in the fields of medicine and health science, technical science, and natural science—which should take precedence over word-by-word phrasing and any attempt to modify the text to keep the similarity index below a certain, arbitrary threshold.

If the focus in evaluating the manuscript’s originality is shifted from the whole and the essence to the pure form and word-by-word phrasing, we wonder whether, in the future, the overly mechanical and administrative handling of text similarity will stimulate the development or adoption of novel dominant formats for presenting scientific results (e.g., graphical and/or video instead of text).

Author Contributions: Conceptualization, P.M. and M.D.; formal analysis, P.M., T.P. and M.D.; writing—original draft preparation, P.M.; writing—review and editing, T.P. and M.D.; supervision, M.D. and T.P.; project administration, T.P.; funding acquisition, M.D. and T.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia (grant number 451-03-66/2024-03/200110 and grant to the Center of Bone Biology) and the Science Fund of the Republic of Serbia (grant number 7749444, BoFraM).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

Akbari, A. (2021). Spinning-translation and the act of plagiarising: How to avoid and resist. *Journal of Further and Higher Education*, *45*, 49–64. [CrossRef]

Alvi, F., Stevenson, M., & Clough, P. (2021). Paraphrase type identification for plagiarism detection using contexts and word embeddings. *International Journal of Educational Technology in Higher Education*, *18*, 42. [CrossRef]

Ansorge, L. (2024). Tortured phrases are not automatically unethical. *European Science Editing*, *50*, e135388. [CrossRef]

Ansorge, L., Ansorgeová, K., & Sixsmith, M. (2021). Plagiarism through paraphrasing tools—The story of one plagiarized text. *Publications*, *9*, 48. [CrossRef]

Aronson, J. (2021, August 6). *When I use a word . . . artificial translation*. Available online: https://blogs.bmj.com/bmj/2021/08/06/jeffrey-aronson-when-i-use-a-word-artificial-translation (accessed on 29 November 2024).

Baskaran, S., Agarwal, A., Panner Selvam, M. K., Henkel, R., Durairajanayagam, D., Leisegang, K., Majzoub, A., Singh, D., & Khalafalla, K. (2019). Is there plagiarism in the most influential publications in the field of andrology? *Andrologia*, *51*, e13405. [CrossRef] [PubMed]

Bretag, T., & Mahmud, S. (2009). Self-plagiarism or appropriate textual re-use? *Journal of Academic Ethics*, *7*, 193–205. [CrossRef]

Cabanac, G., & Labbé, C. (2021). Prevalence of nonsensical algorithmically generated papers in the scientific literature. *Journal of the Association for Information Science and Technology*, *72*, 1461–1476. [CrossRef]

Cabanac, G., Labbé, C., & Magazinov, A. (2021). Tortured phrases: A dubious writing style emerging in science. Evidence of critical issues affecting established journals. *arXiv*, arXiv:2107.06751. [CrossRef]

Devlin, M. (2006). Policy, preparation, and prevention: Proactive minimization of student plagiarism. *Journal of Higher Education Policy and Management*, *28*, 45–58. [CrossRef]

Else, H. (2021). 'Tortured Phrases' give away fabricated research papers. *Nature*, *596*, 328–329. [CrossRef] [PubMed]

Gangadharan, V., Gupta, D., Amritha, L., & Athira, T. A. (2020, June 15–17). *Paraphrase detection using deep neural network based word embedding techniques*. 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI) (48184) (pp. 517–521), Tirunelveli, India.

Gunnarsson, J., Kulesza, W. J., & Pettersson, A. (2014). Teaching international students how to avoid plagiarism: Librarians and faculty in collaboration. *The Journal of Academic Librarianship*, *40*, 413–417. [CrossRef]

Higgins, J. R., Lin, F.-C., & Evans, J. P. (2016). Plagiarism in submitted manuscripts: Incidence, characteristics and optimization of screening—Case study in a major specialty medical journal. *Research Integrity and Peer Review*, *1*, 13. [CrossRef] [PubMed]

Jones, M. (2009, September 28–30). *Back-translation: The latest form of plagiarism*. The 4th Asia Pacific Conference on Educational Integrity (pp. 1–7), Wollongong, Australia.

Jones, M., & Sheridan, L. (2015). Back translation: An emerging sophisticated cyber strategy to subvert advances in 'Digital Age' plagiarism detection and prevention. *Assessment & Evaluation in Higher Education*, *40*, 712–724. [CrossRef]

Kannangara, D. N. (2017). Quality, ethics and plagiarism issues in documents generated using word spinning software. *MIER Journal of Educational Studies Trends and Practices*, *7*, 24–32. [CrossRef]

Lay, P., Lentschat, M., & Labbe, C. (2022). Investigating the detection of tortured phrases in scientific literature. In A. Cohan, G. Feigenblat, D. Freitag, T. Ghosal, D. Herrmannova, P. Knoth, K. Lo, P. Mayr, M. Shmueli-Scheuer, A. de Waard, & L. L. Wang (Eds.), *Third workshop on scholarly document processing* (pp. 32–36). Association for Computational Linguistics.

Mahian, O., Treutwein, M., Estellé, P., Wongwises, S., Wen, D., Lorenzini, G., Dalkilic, A. S., Yan, W.-M., & Sahin, A. Z. (2017). Measurement of similarity in academic contexts. *Publications*, *5*, 18. [CrossRef]

Manley, S. (2023). The use of text-matching software's similarity scores. *Accountability in Research*, *30*, 219–245. [CrossRef]

Martel, E., Lentschat, M., & Labbé, C. (2024). Detection of tortured phrases in scientific literature. *arXiv*, arXiv:2402.03370.

Martínez-López, J. I., Barrón-González, S., & Martínez López, A. (2019). Which are the tools available for scholars? A review of assisting software for authors during peer reviewing process. *Publications*, *7*, 59. [CrossRef]

McCuen, R. H. (2018). Advancing scientific knowledge: Ethical issues in the journal publication process. *Publications*, *6*, 1. [CrossRef]

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*, *35*, e217. [CrossRef]

Milovanović, P., Stolić, D., & Pekmezović, T. (2023). Writing phd thesis in English: Importance, challenges, and thesis originality. *Medicinska Istraživanja*, *56*, 43–48. [CrossRef]

Morán, D., & Carlos, R. (2022). The perception of academic plagiarism in industrial engineering students at a public university in Lima. *Publications*, *10*, 41. [CrossRef]

Mostofa, S. M., Tabassum, M., & Ahmed, S. M. Z. (2021). Researchers' awareness about plagiarism and impact of plagiarism detection tools—Does awareness effect the actions towards preventing plagiarism? *Digital Library Perspectives*, *37*, 257–274. [CrossRef]

O'Grady, C. (2024, May 31). *Software that detects 'tortured acronyms' in research papers could help root out misconduct*. Available online: https://www.science.org/content/article/software-detects-tortured-acronyms-in-research-papers (accessed on 1 November 2024).

Perkins, M., Gezgin, U. B., & Roe, J. (2020). Reducing plagiarism through academic misconduct education. *International Journal for Educational Integrity*, *16*, 3. [CrossRef]

Potter, W. (2024, July 23). *An academic publisher has struck an AI data deal with microsoft—Without their authors' knowledge*. Available online: http://theconversation.com/an-academic-publisher-has-struck-an-ai-data-deal-with-microsoft-without-their-authors-knowledge-235203 (accessed on 29 November 2024).

Pourret, O. (2024). On the emergence of tortured phrases: A threat to scientific integrity—The example of "heavy metal". *European Science Editing*, *50*, e131771. [CrossRef]

Prentice, F. M., & Kinden, C. E. (2018). Paraphrasing tools, language translation tools and plagiarism: An exploratory study. *International Journal for Educational Integrity*, *14*, 11. [CrossRef]

Radikė, M., & Camm, C. F. (2022). Plagiarism in Medical Publishing: Each of Us Can Do Something about It. *European Heart Journal—Case Reports*, *6*, ytac137. [CrossRef]

Teixeira da Silva, J. A. (2023). "Tortured Phrases" in preprints. *Current Medical Research and Opinion*, *39*, 785–787. [CrossRef]

Vrbanec, T., & Meštrović, A. (2020). Corpus-based paraphrase detection experiments and review. *Information*, *11*, 241. [CrossRef]

Vrbanec, T., & Meštrović, A. (2023). Comparison study of unsupervised paraphrase detection: Deep learning—The key for semantic similarity detection. *Expert Systems*, *40*, e13386. [CrossRef]

Wahle, J. P., Ruas, T., Foltýnek, T., Meuschke, N., & Gipp, B. (2022). Identifying machine-paraphrased plagiarism. In M. Smits (Ed.), *Information for a better world: Shaping the global future* (pp. 393–413). Springer International Publishing.