Journal of Informetrics 16 (2022) 101304

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Exploring perception of retraction based on mentioned status in post-retraction citations

Xiaojuan Liu [a], Chenlin Wang [a], Dar-Zen Chen [b], Mu-Hsuan Huang [c,*]

[a] School of Government, Beijing Normal University, No. 19, Xinjiekouwai Street, Haidian, Beijing, 100875, China
[b] Department of Mechanical Engineering, National Taiwan University, Roosevelt Road, No. 1, Sec. 4, Taipei, 10617, Taiwan
[c] Department of Library and Information Science, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617, Taiwan

ARTICLE INFO

Keywords:
Post-retraction citation
Full-text analysis
Unaware duration

ABSTRACT

The way retracted papers have been mentioned in post-retraction citations reflects the perception of the citing authors. The characteristics of post-retraction citations are therefore worth studying to provide insights into the prevention of the citation chain of retracted papers. In this study, full-text analysis is used to compare the distinctions of citation location and citation sentiment—attitudes and dispositions toward the cited work—between the conditions of correctly mentioning the retracted status (called CM) and not mentioning the retracted status (called NM). Statistical test is carried out to explore the effect of CM on post-retraction citations in the field of psychology. It is shown that the citation sentiment of CM is equally distributed as negative, neutral, and positive, while for NM, it is mainly distributed as the latter two. CM papers tend to cite retracted papers in *Methodology*, whereas NM papers cite more in *Theoretical Background* and *Conclusion*. The perception efficiency of retractions in psychology is low, where the average unaware duration (UD, the period between when the retraction note has been published and when the first citation directly pointed out its retracted status) lasts for 2.88 years. Also, UD is negatively correlated with the quantity of CM and the growth rate of NM, the proportionate change of NM before and after the first CM paper appears ($P < 0.01$). After being aware of retractions, the average rate of change (ARC, the total change divided by its taken time) of NM declines significantly ($Z = -2.823$, $P < 0.01$) whereas CM sees a raise in most disciplines, which contributes to the reduction of possible interdisciplinary impact.

1. Introduction

In general, a paper that has already been retracted should not be continually spread or cited as a reference. Papers that were officially announced as retractions by the publisher would no longer exist in academic records (Teixeira & Bornemann-Cimenti, 2017) and should not be considered as scientific researches (Van & Nijveen, 2016). In addition, citing retracted papers not only negatively influences the academic discipline, but it also places negative impacts on research which cited the retracted papers, judging by how a certain number of papers are retracted because they had cited prior retracted works (Retraction Watch, 2018). Some papers are still being frequently cited as valid references despite that they had been retracted (Couzin, 2006; Bar-Ilan & Halevi, 2017). Very often, works that cited retracted papers did not mention anything about the fact that such citations come from retracted papers (Bornemann-Cimenti, Szilagyi & Sandner-Kiesling, 2016; Neale, Dailey & Abrams, 2010). Research shows that it takes a long time to naturally eliminate the impacts brought by retracted papers (Furman, Jensen & Murray, 2012); the average period of papers being used as citations after facing retraction is 4 years (2 years as the median) (Chen et al., 2012).

* Corresponding author.
E-mail address: mhhuang@ntu.edu.tw (M.-H. Huang).

https://doi.org/10.1016/j.joi.2022.101304
Received 19 January 2022; Received in revised form 11 May 2022; Accepted 25 May 2022
Available online 4 June 2022

In 1990, Pfeifer and Snodgrass (1990) pointed out the issue of the "continued use of retracted, invalid scientific literature." The academic term commonly used to refer to the situation is "post-retraction citation" or "continued citation." It means the citations that happened after the release of the retraction note (Bar-Ilan & Halevi, 2017; Hamilton, 2019). However, this definition is vague due to the uncertain period from the citing moment to the publication time. As a result, most scholars consider the paper that cited a retracted paper one year after its retraction as a "post-retraction citation paper" (Bar-Ilan & Halevi, 2017; Pfeifer & Snodgrass, 1990; Budd, Sievert & Schultz, 1998; Kochan & Budd, 2010), because it normally takes 12 months (or longer) for a paper to reach final publishing from its first-time submission, depending on the publication cycle of different journals (Neale et al., 2007); moreover, establishing the retraction note index in the database also takes up some time (Budd, Sievert & Schultz, 1998). Hamilton et al. (2019) provided a more precise definition: any publication that referenced a retracted study for which the date of submission was after the date of retraction notice publication; however, if the submitted date is unavailable, the citing papers published six months after the retracted date can be considered as post-retraction citation papers. We believe that the relationship between citing and cited papers is established formally after the publication. Hence, it is more accurate to define the range of post-retraction based on the publication date directly. In this study, the "post-retraction citation" is defined as any continued citation whereas the publication date is after the publication date of the retraction note of the cited retracted papers.

Providing explanation in the content, or labeling "retracted" in the references, some post-retraction citation papers would be clear in conveying the retracted status of their cited references; while the others mention nothing about the retracted status of their references, which, may be a result of unknown changes of the cited paper for both the authors and publishers. In this study, the two categories are respectively called the Correctly Mentioned (CM) citation and the Not Mentioned (NM) citation. The two actually reflect the different levels of perception about retractions in academia. In this study, we comprehensively consider the characteristics of post-retraction citations and illustrate the differences by using citation analysis methods. By doing so, we aim to provide multilateral perspectives about the perception, management, and prevention of post-retraction citation for researchers, publishers, and policymakers.

Guided by this intention, this study focuses on the perception of post-retraction citation scholars and analyzes the characteristics and differences between CM and NM. The several research questions are proposed as follows:

- How are CM and NM distributed in post-retracted citations?
- For CM, in which part of the article do the authors mention the retraction? Is it the main body, the footnote, or simply the reference?
- How do reasons and behaviors impact the post-retraction citation? In what circumstances would CM most likely happen in a post-retraction citation?
- What are the differences between CM and NM about citation location, citation sentiment, and citation intensity?
- What is the influence of CM on retracted papers? After the first appearance of CM, does it decrease the continued citations, the amount of NM, or the growth rate of NM (the proportionate change of NM before and after CM's first appearance)? And to what extent does it decrease?
- On the premise of the appearance of CM, how do scholars from other disciplines and journals view the retractions that occurred in the psychological field? How does it impact the continued citation behavior?

2. Related work

Many scholars have studied the retraction phenomenon, especially the reasons for retractions. The reasons behind the retraction imply different aspects and levels of negative effect. The retracted papers, besides having questionable research outcomes, directly or indirectly influence the accuracy of other academic productions and impact the career of authors concerned. They may even eventually lead to distrust between scientific research and society. By encoding the retraction notice or applying meta-analysis on the present literature, scholars have concluded some reasons for retractions. Nath et al. (2006) divided the reasons into two broad categories: misconduct and unintentional error. Based on the research on retraction notices and retracted papers, Retraction Watch (2018) summarized the reasons for retraction into 94 labels, and further explained each of them. Driven by the editorial guidelines of COPE (Wager et al., 2009), Tian et al. (2014) studied the present literature with the meta-analysis method and divided the reasons for retractions into nine categories, which include fake data, error/mistake, could not replicate findings, plagiarism, duplication, author dispute, lack of consent/approval, publisher error, and others.

On the basis of clarifying the reasons for retraction, scholars go further to study the aftermath impacts brought by the retraction, that is, the continued citation of the retracted papers. Current related research mostly focuses on the cause of post-retraction citation, certain specific academic misconduct (Bornemann-Cimenti, Szilagyi & Sandner-Kiesling, 2016; Kochan & Budd, 2010), the features of post-retraction citation in a specific field (Hamilton, 2019) in terms of its frequency, and the citation sentiment of post-retraction citation in the citing paper.

Scholars divided the cause of post-retraction citation into four types: (1) The journal/publisher website or database failed to update the retracted status or to conduct various update policies on time (Wright & Mcdaid, 2011; Suelzer et al., 2019). (2) Researchers kept the original version or printed copies without tracking the latest status of references (Teixeira & Bornemann-Cimenti, 2017; Couzin, 2006). (3) Multiple copies of the original paper online would no longer be updated (Teixeira & Bornemann-Cimenti, 2017; Rgac, Ag & Mka, 2019). (4) Although being retracted, the methodology, discovery, and conclusion of the papers remain authentic and effective (Teixeira & Dobránszki, 2017). The former three conditions lead to the outcome that the authors are unaware that they are citing retracted papers. As for the fourth condition, the authors usually are aware of the retracted status of their references.

Table 1
Classifications of Retraction Reasons.

Reasons for Retraction	Practical Types	Description
Intentional Misconduct	Plagiarism	The paper copies, imitates, uses others' charts and data improperly or without permission, and does not include relevant citations.
	Fake Data	The paper forges data, alters data results based on subjective views, distorts truths in the conclusion, or makes false claims.
	Duplication	The paper has a whole or partial duplicate publication with the content, data charts, or conclusion.
Unintentional Error	Could Not Replicate Findings	Even under identical experimental conditions, with identical equipment and methodology, the author(s), committee, or a third party are unable to reach identical results. The experimental results cannot be reproduced.
	Error/Mistake	The paper has an inaccurate calculating process or outcome, and there is no evidence of proving it as intentional forgery or deception, which can also be referred to as an "honest mistake."
Other Reasons	Publisher Error	Errors are simply due to mistakes made by the publisher and are not related to the author(s). Such as premature publication.
	Author Dispute	Retractions due to the author(s)'s controversies over signatures, contributions, or permission.
	Ethical Issues	The paper has ethical conflicts, unethical issues, or is without the consent of experimental subjects.
	No reason/unclear	No reasons are given in the retraction note.

They are likely to agree with some perspectives mentioned in the retracted papers, or they would like to study some phenomena caused by the retracted works. On that note, Kim et al. (2019) found that papers that were retracted due to duplicate publication still have a scientific effect. Accordingly, they still got positive citations. From Bar-Il An et al. (2017), if the retracted papers were getting significant social attention from the public and media, they would raise the number of post-retraction citations, including both positive citations and negative citations.

Studies on the characteristics of post-retraction citation frequency analyze the features and trends of the citations as they develop along with time after being retracted. Budd et al. (1999) researched on the post-retraction citation in Biomedical Science and found that 235 papers retracted from 1966 to 1996 were cited 2034 times in the field until 1997. Among them, 275 citing papers showed obvious affirmation to the retracted papers. Neale et al. (2007) studied 102 retracted papers and pointed out that most of these papers were still routinely cited. Scholars who cited them were unaware of the works' nature of being influenced by academic misconduct. Bar-Il An et al. (2017) conducted case studies on 15 retracted papers and their post-retracted citation condition. The research found that even though the publisher had released a retraction notice, the retracted papers were still cited in general, with some of them being cited even more than before.

The citation sentiment toward retracted papers can directly reflect the scholars' attitude toward the retraction works that they cited and also indirectly imply their perception of these retractions. In the current researches, there are sufficient citation sentiment studies on non-retracted papers. The general categories of sentiment division are positive, negative (Moravcsik & Murugesan, 1975; Chubin & Moitra, 1975), neutral (Liu et al., 2014), and mixed (Zhang, Ding & Milojevi, 2013). The sentiment studies on post-retraction citation of retracted papers are based on these categories and most scholars conduct full-text analysis with the empirical data. Bar-Il An et al. (2017) divided the sentiment into three categories as positive, negative, and neutral, viewing through the context of the citation with text analysis method and showing that most of the post-retraction citations are positive. Wright (1991) divided the sentiment into two types as affirmative and negative. He indicated that the negative citation is mainly a reminder for readers about the errors in the retracted papers. Korpela (2010) found that the affirmative citing of fraudulent research had not completely stopped but continued for 24 years after the retraction in the Breuning case. Hamilton (2019) divided the post-retraction citation sentiment into positive and negative, and he pointed out the positive means to cite the retracted papers as normal ones while the negative intends to point out the retracted status while citing the reference. For Catalini et al. (2015), their research showed that negative citations are most likely to appear in the section of "discussion" and "conclusion." In such cases, the author of the citing paper usually points out the inconsistent outcome of the research with the cited one.

3. Methods

In response to the research questions raised, this study applies full-text analysis to compare differences between CM and NM, encoding citation characteristics such as: reasons for retraction, citation location, citation sentiment, and citation intensity as indicators. The clue words with high frequency are analyzed statistically to understand the content components. According to these characteristics, new quantitative indicators —the unaware duration and the changing of NM— are proposed to measure the scholars' perception of the retracted papers. As for the causes of retraction, the categories summarized by Nath et al. (2006) and Tian et al. (2014) are integrated into three broad categories and nine practical types, as shown in Table 1. Through classifying the retraction notices by manual means, we obtained the reasons for the retraction of each retracted paper. For cases that have more than one reason, they are counted as many times as the number of reasons.

3.1. Full-text analysis

Taking the full text of citing paper as the subject for study, the full-text analysis conducts data mining, quantitating, and analyzing to information such as citation location, citation sentiment, and citation intensity (Zhao, Zeng & Chen, 2014). The research encodes

Table 2
Classifications of post-retraction citation locations.

Location	Actual Section Headers
Abstract	Abstract; Author summary; Summary
Introduction	Introduction; Background; First three paragraphs (without title)
Theoretical Background and Hypothesis	Literature review; Conceptual background; Theoretical background; Theoretical development and hypotheses; Literature review and hypotheses; Theoretical framework; Related work; Theoretical basis; Research background……
Case Study	Case Study; Experiment(s); Study overview
Methodology	Method(s); Material(s) and method(s); Material(s) & method(s); Methodology; Models and methods; Data; Dataset; Subjects and methods; Experimental procedures……
Discussion and Conclusion	Data analysis; Statistical analysis; Result(s); Finding(s); Result(s) and discussion(s); Result(s) & discussion(s); Discussion(s); Conclusion(s); Discussion(s) and conclusion(s); General discussion; Recommendations; Future Research; Summary (at the end of the text)……
Appendix and Notes	Note(s); Appendix; Footnote(s); Further reading(s)
References	No citation in the main text but only listed in the references

Table 3
Standards for classifying citation sentiment of retracted papers.

Sentiment	Reference Purpose	Description
Positive	Basic	The citing paper's research foundation and research results are based on the retracted paper. For example, the basic equation, experimental conditions, evaluation table, experimental hypothesis, and data are referenced or came from the retracted papers.
	Supplementary	The independent views or insights from the retracted papers are cited as supportive ideas in the citing paper.
Neutral	Perfunctory	The retracted papers are mentioned in the citing papers merely as related literature without further critique or description, or they are just listed in the reference.
Negative	Partial Negational	The citing papers point out the mistakes/errors of the cited ones or supplement their insufficient parts.
	Total Negational	The citing papers regard the cited ones as wrong and provide them independent solutions.

the data from manual judgment to grab the cited parts and locations, and to integrate with the context to obtain the sentiment of citation. On this basis, citation intensity and high frequency clue words are statistically analyzed to figure out the characteristics of citations from the aspect of the content.

3.1.1. Recognizing citation location

This article refers to research on the general structure (Luong, Nguyen & Kan, 2010) and citation location (An et al., 2017; Peng, Yan & Huang, 2019) of academic documents, and marks the location of the citation fragments of the cited papers after the retraction. The specific location categories are shown in Table 2.

3.1.2. Recognizing citation sentiment

The analysis of the citation sentiment is one of the crucial parts of full-text analysis, and the rationality of the encoding content directly influences the accuracy of the outcome. Based on a rigorous, scientific attitude, this research aims to combine the current categories of citation sentiment studies (Bar-Ilan & Halevi, 2017; Chubin & Moitra, 1975) with characteristics of retractions, and divide the sentiments into three broad categories, with each category having a different specific reference purpose, as shown in Table 3.

3.1.3. Citation intensity statistics

An academic paper might be cited several times within one citing work. Some references are mentioned several times within the full text of a paper and others might be mentioned only once (Pak, Wang & Yu, 2020). For a reference of the paper, the number of times it has been mentioned in the article represents the citation intensity (Zhang, Liu & Wang, 2021).To be more specific, when a cited paper is mentioned once in the citing paper, the citation intensity would be counted as 1, if mentioned for several times, the number goes up. Regarding the cross-statistics of citation location and citation intensity, the citation intensity of CM and NM in different locations are acquired; then the classified statistics of CM and NM are divided into three sentiment categories (positive, neutral, and negative), and finally, the citation intensity index of the different citation sentiments can be obtained. Therefore, the citation intensity of a certain location or sentiment and its proportion can be calculated as follows:

$$\text{citation intensity(i)} = n_i^{(j)}$$

$$\text{proportion of citation intensity(i)} = \frac{\sum_j n_i^{(j)}}{\sum_j \sum_i n_i^{(j)}} \quad \text{(I)}$$

In *Formula I*, $n_i^{(j)}$ is the number of times that these references are mentioned in location i (or sentiment i) of article j, while $\sum_i n_i^{(j)}$ means the number of times of these references appearing in all locations (or sentiments) of article j.

3.1.4. Clue words extraction

Teufel (1999) studied the rules of classification on citation context in his doctoral dissertation, and manually selected 892 clue words, constructing the most comprehensive clue words dictionary for citation studies. Based on this dictionary, this study uses NLTK to sort out non-denotative words such as study, work, and research; and also subject-related terms such as ketamine, systematic, or school shooting. Ultimately, the 10 clue words with the highest frequency can be acquired, and be integrated with citation location and citation sentiments for in-depth analysis.

3.2. Statistical testing method

3.2.1. Calculating unaware duration

A retracted paper should theoretically be acknowledged by other scholars when the retraction notice has been released, but in the actual condition, the delay of perception usually happens, causing some cases of NM citations. Some studies have already shown that owing to the unawareness of citing retractions, NM papers cite without further explanation of the sources' retracted status (Bornemann-Cimenti, Szilagyi & Sandner-Kiesling, 2016; Neale, Dailey & Abrams, 2010; Neale et al., 2007). These papers have misled more scholars into believing that the references of NM papers are non-retracted and therefore resulted in the continued citations of the retracted papers, thereby strengthening the citation chain of retracted papers. By then scholars would still remain unaware of the issue (Weinstein, Sandman & Blalock, 2008), as well as the consequences that followed. Therefore, in the citation chain of retracted papers, if a citing paper first specifically mentioned the retracted status of a cited paper, it means that the communities besides the stakeholders of the retracted paper - including the authors, affiliated journals, and the publishers - also recognize the retracted status of the paper, and this paper will then be marked as CM0. In this study, the interval between the time when a paper is retracted and the publication time of CM0 is defined as the Unaware Duration (UD). Worth noting, for those retracted papers without CM citations, all the papers in its citation chain are NM papers. Under this circumstance, the UD is an infinity (∞). The UD of retracted papers is expressed as the formula below:

$$\mathrm{UD} = \begin{cases} T_{CM0} - T_{retract}, & \exists P_{CM} \in P \\ \infty \ , & \nexists P_{CM} \in P \end{cases} \tag{II}$$

In *Formula II*, P = {P1, P2,...,Pn} refers to the collection of post-retraction citation papers of a retracted paper; P_{CM} refers to CM papers, and T_{CM0} refers to the released date of the first CM paper; $T_{retract}$ is the publication date of the retraction note.

3.2.2. Analyzing the change of unawareness

Around the time when CM0 appears, the variation of continued citations, NM's growth rate, and NM's average rate of change can indicate the practical influence of CM papers, through which whether the "correctly mentioned citation" could enhance other scholars to notice the retractions on journals or papers can be further explored.

NM's growth rate, which means the proportionate change of a specific variable within a specific period, is noted as ΔNM' and the calculation is as follows:

$$\Delta\mathrm{NM}' = \mathrm{NM}'\mathrm{s\ growth\ rate} = \frac{NM^{(T_i)} - NM^{(T_{i-1})}}{NM^{(T_{i-1})}} \tag{III}$$

In *Formula III*, T_i is the time when CM0 appears. $NM^{(T_{i-1})}$is the number of NM before CM0, $NM^{(T_i)}$is the number of NM after CM0.

The average rate of change (ARC) of NM means the changing value of quantity during the elapsed time, which is the same as the derivative of the NM(T) function, the slope of its curve. From the period of being retracted to CM0, also known as UD of retracted papers, the ARC of NM is noted as $k_{\Delta\mathrm{NM0}}$. Also, recording the ARC of NM during the period from CM0 to the last NM (NMt) in the sample as $k_{\Delta\mathrm{NMt}}$. The basic formula of ARC is:

$$k_{\Delta\mathrm{NM}} = \mathrm{NM}'\mathrm{s\ Average\ Rate\ of\ Change} = \frac{\Delta NM}{\Delta t} = \frac{NM^{(T_i)} - NM^{(T_{i-1})}}{T_i - T_{i-1}} \tag{IV}$$

In *Formula IV*, $NM^{(T_i)}$ is the number of NM in T_i. Based on this formula, $k_{\Delta\mathrm{NM0}}$ and $k_{\Delta\mathrm{NMt}}$ can be calculated as follows:

$$k_{\Delta\mathrm{NM0}} = \frac{NM^{(T_{CM0})} - NM^{(T_{retract})}}{T_{CM0} - T_{retract}} = \frac{Number\ of\ NM\ before\ CM0}{UD} \tag{V}$$

$$k_{\Delta\mathrm{NMt}} == \frac{NM^{(T_{NMt})} - NM^{(T_{CM0})}}{T_{NMt} - T_{CM0}} = \frac{Number\ of\ NM\ after\ CM0}{T_{NMt} - T_{CM0}}, \tag{VI}$$

where $NM^{(T_{NMt})} - NM^{(T_{CM0})} \neq 0$.

Particularly, when $NM^{(T_{NMt})} - NM^{(T_{CM0})} = 0,\ k_{\Delta\mathrm{NMt}} = 0$.

Statistical and correlation analysis has been performed on the data in Excel and SPSS. Furthermore, the Spearman test is used to analyze the correlation of the three variables of non-normal distribution, that is, UD, NM's growth rate (ΔNM'), and the continued citation numbers after CM0. With Wilcoxon's Sign Rank Test, the ARC of NM before and after CM0 ($k_{\Delta\mathrm{NM0}}$, $k_{\Delta\mathrm{NMt}}$) are compared to figure out how the condition of "citing papers correctly mentioned about the retracted status" will affect the post-retraction citations.

Table 4
Sample dataset.

Object	Attribute	Quantity
Retracted Paper	Author, title, online published date, official publication date, abstract	93
Retraction Note	Author, title, retracted date, retraction reason, retraction behavior	93
Post-retraction Citing Paper	Author, title, subject, online published date, official publication date, citation fragment, mentioned status, citation location, citation sentiment, citation purpose, citation intensity	1010

4. Data

4.1. Search strategy

The object of this research is the post-retraction citing papers, therefore, all the relevant metadata or information including retracted papers, retraction notes, and original text of cited papers. The process of data collecting is as follows:

First, with certain standards (e.g. Essential Science Indicators (ESI) journal list contains 22 broad disciplines and more than 11000 journals from Web of Science Core Collection database, in which a journal is assigned to only one discipline) the retracted papers are retrieved. Then, perform an advanced search in Web of Science according to the search formula TI=Retracted article, and select the document type as “retracted” or “retracted publication” to obtain all the retractions within a certain period in a specific field. The original texts of the retracted papers are collected by recording the basic information about authors, titles, subjects, etc., and time-related crucial information such as the online published date and official publication date are recorded as well.

Secondly, download the retraction notes based on the information of retracted papers from the publishers’ website, acquire the retracted date, the retraction reason, and the text. After that, go further to encode information for analysis.

Finally, obtain the citing papers. According to the database of retracted papers in Web of Science, each of their citing papers is downloaded and their basic information, authors, titles, subjects, online published date, and official publication date are recorded as metadata. The sifting of these post-retraction citing papers is based on the definition of post-retraction citation mentioned in this paper’s *Introduction*. According to the post-retraction citing papers’ mentioning of the retracted papers, they can be divided into two categories, CM and NM.

4.2. Case

Current research on retractions mostly centers around the biomedicine field (Bar-Ilan & Halevi, 2017), yet empirical studies in social sciences and interdisciplinary fields still need further exploring. Since psychology is an interdisciplinary discipline integrating nature and social science, the case analysis of post-retraction citations in the field could be representative. The characteristics of post-retraction citations in the field also bring more concrete research processes and outcomes.

The empirical objects of this study are retracted papers, retraction notes, and post-retraction citing papers in the field of *Psychiatry/Psychology*. According to journals of *Psychiatry/Psychology* in the latest ESI journal list (ESI-2021), our data was acquired from Web of Science databases and all data was collected in March 2021. There are 120 searching results of retracted papers with retraction notes before December 31, 2020. With the 120 retracted papers, there are 2493 citing papers in total, including 1045 post-retraction citing papers which cover 93 retracted papers. Among these post-retraction citing papers, 1010 of them have the full-text available for access (the proportion is 96.65%). Within these papers, the cited parts are extracted, encoded, and classified to obtain information such as their mentioned status, citation location, citation sentiment, citation purpose, and citation intensity (Table 4).

5. Results and discussion

5.1. Overview of post-retraction citations

Among the 1010 post-retraction citing papers, only 108 of them correctly mentioned the retracted status of the reference which includes three locations: (1) 6.04% only in the reference section; (2) 3.96% in the text; (3) 0.69% in the form of notes. The cases mentioned above are classified as Correctly Mentioned (CM) citations, while the others belong to Not Mentioned (NM) citations (Table 5).

According to Table 1, the corresponding retraction notes from 93 post-retraction citing papers are encoded and statisticized, while 2 of them have two retraction reasons. As shown in Table 6, retractions due to misconduct take up the proportion of 60%. In the statistics of practical reasons, fake data takes up 49.47% followed by error/mistake with 32.63%. From the analysis, the post-retraction citation number of misconduct papers is exceeding the unintentional papers and their CM citations are equal to the latter. Specifically, CM papers mainly cite the retracted papers with fake data and errors; as for papers retracted for other reasons such as author dispute, publisher errors, and ethical issues, they have CM citations.

For the retractions, most of the passive ones are initiated by editors after the journals, publishers, institutions, or other scholars have discovered the misconduct of the papers. Voluntary retractions, on the contrary, are requested spontaneously by authors. In Table 7, papers retracted due to misconduct are mainly passive retractions, while the types of unintentional errors are more likely

Table 5
Mentioned status distribution of post-retraction citation.

Mentioned Status	Mentioned Position	Count	Proportion
CM	Only in the references	61	6.04%
	In the text	40	3.96%
	In the notes	7	0.69%
	Total	108	10.69%
NM	-	902	89.31%
Sum		1010	

Table 6
Post-retraction citations of different retraction reasons.

Retraction Reason	Specific Reason	Retracted Paper	Citation	CM	NM
Intentional Misconduct	Plagiarism	8	141	5	136
	Fake Data	47	355	53	302
	Duplication	2	52	0	52
	Total	57	548	58	490
Unintentional Error	Could Not Replicate Findings	1	46	10	36
	Error/Mistake	31	449	48	401
	Total	32	495	58	437
Other Reasons	Publisher Error	2	2	1	1
	Author Dispute	1	8	3	5
	Ethical Issues	1	3	1	2
	No reason/unclear	2	8	0	8
	Total	6	21	5	16

Table 7
Post-retraction citations of different retraction behaviors.

Retraction Behavior	Quantity of Retracted Papers	Retraction Reason (Horizontally)			Post-retraction Citation					
		Intentional Misconduct	Unintentional Error	Other Reason	Overall		CM		NM	
					Sum	AVG	Sum	AVG	Sum	AVG
Passive Retraction	65	76.92%	18.46%	4.62%	555	8.54	47	0.72	508	7.82
Voluntary Retraction	28	23.33%	66.67%	10.00%	455	16.25	61	2.18	394	14.07

to be errors found by the authors, and based on which they request for retractions themselves. In sum, the average post-retraction citations, the average number of CM and NM citations of voluntary retractions are higher than the passive ones.

5.2. Content characteristics of post-retraction citations

The post-retraction citations of the 1010 citing papers are 1431 times and the average citations per paper are 1.417 times (1431 times/1010 papers). The citation intensity of the 1010 citing papers is 1.80 (194 times/108 papers) for CM and 1.37 (1237 times/902 papers) for NM. Content features such as the citation location and citation sentiment of CM and NM papers are explained by statistical analysis and word frequency analysis.

5.2.1. Analysis of citation location distribution

To compare the distribution of the citation locations of CM and NM, the most common citation locations of both are *Theoretical Background and Hypothesis*. Relatively, CM papers tend to raise questions from viewpoints of retracted papers in *Introduction*, to use the reference for *Methodology*, the experimental scales, or data acquisition methods of retracted papers, and to explain the reasons for citing the retracted papers in *Appendix and Notes*. The clue words in *Introduction* of CM are “previous,” “question,” etc., while in *Case Study* the words are “contrast,” “reasonable,” etc.; in *Methodology* are “trust,” “standard,” etc., and in *Appendix and Notes* are “believe,” “positive,” etc. From the analysis on keywords with high frequency, the authors of CM papers trust the validity of the research process and methods from the retracted papers. Even if the references were retracted due to their use of fake data or containing errors, parts of their experimental methods still have a certain value. Therefore, even though it is a known fact that the papers have been retracted, there are still CM citations.

In contrast, NM citations appear mostly in *Theoretical Background and Hypothesis* and *Discussion and Conclusion*. On the one hand, retracted papers are used as a theoretical basis to propose research hypotheses; and on the other hand, conclusions are contrasted to support themselves. With the unawareness of the retracted status of reference, NM’s characteristics are similar to general citations. It is worth noting that in *Discussion and Conclusion*, the inconsistency between NM and the retracted paper is sometimes compared to highlight the uniqueness of its research, in which the clue words are “previous,” “increase,” “loss,” etc (Fig. 1).

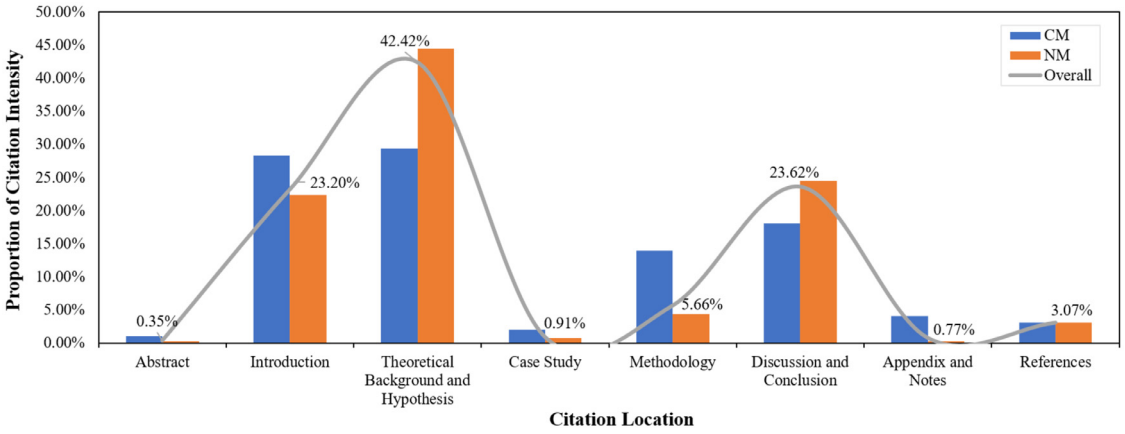

Fig. 1. Citation location distribution of CM vs NM.
*The vertical axis is the ratio of the citation intensity of one location to that of CM and NM respectively.

Table 8
Citation sentiment distribution of different retraction reasons.

Citation Sentiment Retraction Reason		Positive	Neutral	Negative
Intentional Misconduct	Plagiarism	37	159	4
	Fake Data	144	274	51
	Duplication	25	59	0
	Total	206	492	55
Unintentional Error	Could Not Replicate Findings	15	32	6
	Error/Mistake	220	409	31
	Total	235	441	37
Other Reasons	Publisher Error	2	1	0
	Author Dispute	12	18	0
	Ethical Issues	1	2	0
	No reason/unclear	5	7	0
	Total	20	28	0
Sum		461	961	92

5.2.2. Analysis of citation sentiment distribution

Table 8 illustrates the analysis of the post-retraction citation sentiment with different retraction reasons. The comparison shows that whether it is misconduct, unintentional error, or other reasons, the citation sentiment mostly concentrated on neutral citations, and was then followed by positive citations and negative citations.

With the analysis on citation sentiment and the citation intensity as shown in Fig. 2, the proportion of perfunctory citations intensity is up to 64%; the supplementary citations surpass 20% of the total, only 2% is the total negational citations and less than 4% is the partial negational citations. The clue words of the sentiment dimension have more prominent traits. For positive citations, the words mainly include “adapt” and “help,” for neutral ones the words involve “development” and “process,” and for the negative ones they are “contrast,” “how,” and “different,” etc. Concerning the retraction reasons, in positive sentiment, the probability of citing unintentional error papers and intentional misconduct papers is very close, accounting for 14.4% and 14.5% respectively; whereas in negative sentiment, intentional misconduct accounts for 2% more than the unintentional error.

In Fig. 2, the comparison between CM and NM shows obvious distinctions in citation sentiment. NM citations are mainly neutral and positive, while CM citations are more evenly distributed in the three categories. Compared to NM, although CM holds a more distinguishable negative attitude towards retracted paper, its percentage of positive citations still takes out a certain proportion. Then, why do CM papers continue to cite the retracted reference? To further figure out the citation motivations of CM, the citation purposes in different citation locations are counted, and the stacked chart is drawn as Fig. 3.

Figure 3 indicates that CM citations are shown in *Methodology, Case Study*, and *Introduction* as the fundamental basis of the research. The citing authors consider that the scales, methods, and viewpoints from the cited papers are convincing and unaffected by the retraction. The retracted papers can support the scales, methods, hypothesis, conclusions, or viewpoints of the citing papers and even play a key role in their studies as in *Example 1*. All the citation locations, except for the *References*, have supplementary citations which have the same purpose as basic citations. They all present certain consent on the viewpoints of the retracted papers, while the difference is that the former provides weaker support than the latter.

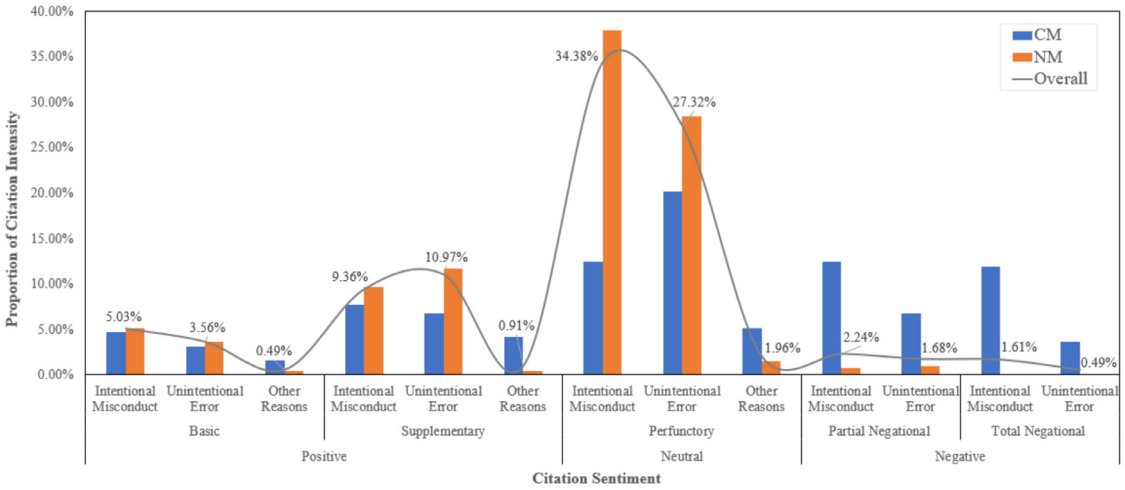

Fig. 2. Citation sentiment distribution of CM vs NM.

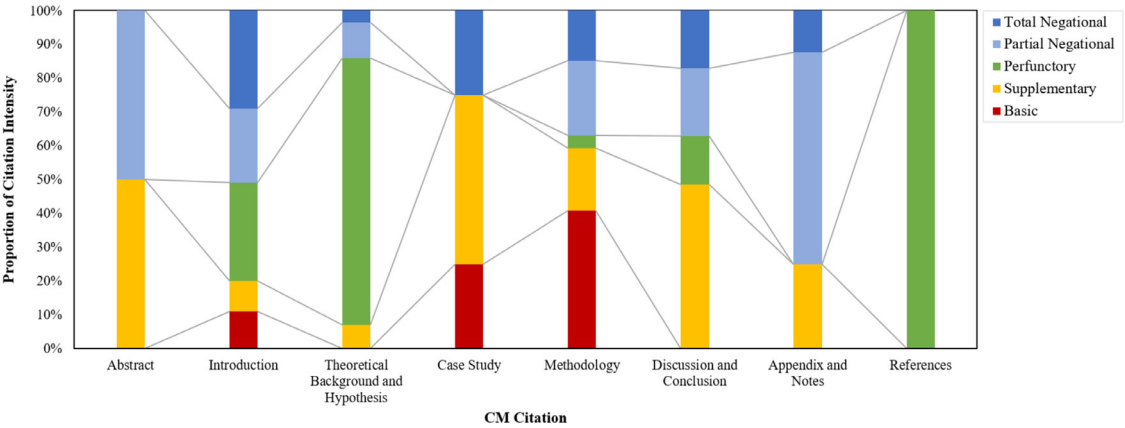

Fig. 3. Stacked bar of citation purpose distribution in different locations of CM.

"Nevertheless, the hypothesis they put forward is a reasonable one and thus we report Experiment 7 in relation to that hypothesis. Note that another report on intelligence priming has also recently been retracted."(Example 1).

Except for *Case Study* and *References*, the remaining six locations are partial negative citations in CM papers. In the part of Abstract, the retraction is usually touched upon as a background to illustrate the research questions, for instance, the citing papers aim to study the uncertain conclusions due to the retraction (e.g. *Example 2*). Besides the purpose mentioned above, in *Introduction, Theoretical Background and Hypothesis*, and *Discussion and Conclusion*, CM papers also tend to fill the research gap of the retracted papers and such gap is generally an unknown part of the themes (e.g. *Example 3*), having no relation with the retraction. This is similar to the partial negative citation in NM papers. In *Methodology* and *Appendix and Notes*, it is shown that after carrying out a certain amount of work, parts of the data are eliminated or needed further verifying since the retracted status of the reference has been discovered. Moreover, CM has total negational citations in six locations including *Introduction, Theoretical Background and Hypothesis*, and *Case Study*. The purposes of these total negational citations involve pointing out the unreliable result caused by misconduct, explaining the elimination of them as samples of the meta-analysis, and dissecting them as cases of academic misconduct.

"However, this report (Stapel & Semin, 2007) was compromised by data fabrication by the first author, so that it remains unclear whether or not linguistic category priming influences perceptual processing. To fill this gap in the literature, the present article reports 12 studies…"(Example 2).

"As discussed in the literature section, authentic leadership has attracted researchers' attention due to its positive role in (Walumbwa, Luthans, Avey & Oke, 2011), calling for more empirical work. We filled this gap by investigating the impact of authentic leadership on communal relationships among employees working in the banking sector of Pakistan." (Example 3).

Table 9
Results of Spearman's Rank-Order Correlation Test.

			UD	CM	ΔNM'
Spearman's rho	UD	Correlation Coefficient	1.000	-.428**	-.705**
		Sig. (2-tailed)	.	.008	.000
		N	37	37	31
	CM	Correlation Coefficient	-.428**	1.000	.425*
		Sig. (2-tailed)	.008	.	.017
		N	37	37	31
	ΔNM'	Correlation Coefficient	-.705**	.425*	1.000
		Sig. (2-tailed)	.000	.017	.
		N	31	31	31

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

5.2.3. Summary of analysis on content characteristics

According to Fig. 2, 64% of the post-retraction citations are neutral, 30% are positive, and merely 6% are negative. This is inconsistent with the results of Bar-Il An et al. (2017) and Suelzer et al. (2019). Bar-Il An et al. (2017) found that the continued positive citations of retracted articles in Elsevier took up 83%, while neutral and negative ones took up 12% and 5% respectively. Suelzer et al. (2019) studied all the citing papers of one retracted work and figured out that negative citations took the highest rate up to 72.7%, while neutral and positive citations took up 9.2% and 8.2% respectively, which were much lower compared to negative citations. From the discrepancy in research outcomes, it is implied that the features of post-retraction citations in diverse disciplines and platforms are distinct, and there are also great differences between individual cases and overall circumstances. Consequently, regarding the management of retracted papers, each discipline and platform must track and mark the cited characteristics of their post-retraction, as well as their negative dissemination. For specific cases which have a wider range of impacts, various ways of management are needed depending on the citation purposes. That is, if there are mostly positive or neutral citations, it is necessary to keep tracking the retracted papers' transmission route and prevent its spread.

Through the analysis of the citation characteristics, it is found that: (1) CM citations clearly distinguish the citation sentiments and purposes. For the location, there are more positive ones in *Methodology*, especially the basic citations. It indicates that the citing scholars agree that the research methods or techniques adopted in retracted papers are still valid and discussable even if these papers have been retracted for other problematic issues. Also, they believe that the enlightenment provided by these papers may heighten the utilization of scientific research products and promote the construction of a research method system. (2) NM citations quote retracted studies as non-retracted papers. The citation locations concentrated in *Introduction, Theoretical Background and Hypothesis*, and *Discussion and Conclusion*. The purposes of citations are primarily perfunctory and supplementary. To sum up, CM citations have a more thorough understanding of both strengths and weaknesses of their retracted references, and the sentiment is more explicit, making more efficient use of the scientific research outcomes. In contrast, the attitude of NM towards the cited papers is more neutral and positive, attributing to the unawareness of retractions and beliefs that these works are regular ones that have been scientifically reviewed in compliance.

5.3. Perception of retraction

The indicators, including unaware duration (UD), the citation numbers before and after CM0, NM's growth rate (ΔNM'), and the ARC of NM before and after CM0 ($k_{\Delta NM0}$, $k_{\Delta NMt}$), are calculated to quantify the awareness of the academic community on the retractions. UD indicates the delayed levels of perception in the time dimension. Citation numbers before and after CM0, ΔNM', $k_{\Delta NM0}$, and $k_{\Delta NMt}$ reveal the influence on awareness of other scholars by correctly mentioning the retracted status. There are two implications by discussing the post-retraction citation from the view of UD and CM0: from the aspect of journal editors and citing authors, doing so can strengthen the examination of references as well as the tracking of retractions; from the aspect of a general scientific research system, it can be utilized to label the retraction within a specific time, and backtrack and block dissemination of retracted works. From the two aspects, both sides work together to intervene in a prolonged trend of unaware duration and also break the citation chain of retracted papers.

5.3.1. Overall perception and changes

Statistics on the data show that 37 of 93 retracted papers are correctly mentioned and these papers have a number of 600 post-retraction citing articles, with a citation intensity up to 882. For the 37 retracted papers, CM0 of each one is acquired and their UD and ΔNM' are computed by *Formula II* and *Formula III* respectively. The results indicate that, the average UD of the samples is 1052.81 days, meaning around 2.88 years. The citation numbers after CM0 rise by an average of 5.43 than before, and the CM citations go up by an average of 1.92. Moreover, the average ΔNM' is +0.63, which represents the growth of unaware citations.

Spearman's Rank-Order Correlation Test demonstrates the relation among UD, the quantity of CM, and ΔNM' (shown in Table 9). As a result, UD and CM numbers are negatively correlated (r=-0.428), which is significant (p=0.008<0.05). Similarly, UD has a strong negative correlation with ΔNM' (r=-0.705, p=0.000<0.01). It is implied that in the field of psychology, the shorter the unaware duration of retraction, the earlier CM0 appears, and there would be more CM citations. For NM citations, the quantity change of each unit would be greater.

Table 10
Descriptive Statistics of Wilcoxon Signed-Rank Test.

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
$k_{\Delta NM0}$	37	.0107	.0174	.0000	.0789	.0011	.0029	.0106
$k_{\Delta NMt}$	37	.0053	.0080	.0000	.0301	.0000	.0015	.0070

Table 11
Ranks and Test Statistics of Wilcoxon Signed-Rank Test.

		N	Mean Rank	Sum of Ranks	Z	Sig.(2-tailed)
$k_{\Delta NMt}$ - $k_{\Delta NM0}$	Negative Ranks	26[a]	16.87	438.50	-2.823	.005
	Positive Ranks	7[b]	17.50	122.50		
	Ties	4[c]				
	Total	37				

[a] $k_{\Delta NMt} < k_{\Delta NM0}$.
[b] $k_{\Delta NMt} > k_{\Delta NM0}$.
[c] $k_{\Delta NMt} = k_{\Delta NM0}$.

Based on *Formula V* and *Formula VI*, $k_{\Delta NM0}$ and $k_{\Delta NMt}$ are calculated respectively, and the difference between them is analyzed through the non-parametric Wilcoxon Signed-Rank Test. The mean value of the ARC of NM before CM0 is 0.0107 and 0.0053 after CM0 (shown in Table 10). Comparing the two values, it can be seen that the growing speed of NM drastically decreases after CM0. The Z value is -2.823, p=0.005<0.01 (shown in Table 11), bearing out the statistical implication that the two variables are significantly different. Compared with the mean values of $k_{\Delta NM0}$ and $k_{\Delta NMt}$, it is certified that the appearance of CM0 has an inhibitory effect on the rising rate of NM of retracted papers.

The results mentioned above indicate that for scholars in the field of psychology, the perception of retraction is low. On the one hand, there is a delayed time from the retracted time to when the scholars recognize the retraction, taking an average of 1052.81 days (approximately 2.88 years) to detect the fact for the first time. On the other hand, the recognized diffusion after the paper has been retracted is poor. It can be seen that the post-retraction citations, especially NM, continue to increase even after the emergence of CM0. But fortunately, a shorter unaware duration signifies an earlier appearance of CM0 and a larger number of CM. In addition, the ARC of NM after CM0 notably decreases compared to that before CM0, which illustrates that the appearance of CM0 can restrain NM's rates of increase.

5.3.2. Discipline perception and changes

Retracted papers in the field of psychology may be cited by journals or papers from other disciplines, causing an interdisciplinary spread and raising a broader range of negative impacts. This study analyzes the distribution of post-retraction citing disciplines of the 37 retracted papers which have CM citations. Furthermore, ESI-2021 is taken as a reference to classify the disciplines that the citing paper belongs to. On this basis, the continuous citations before and after CM0 in diverse disciplines are further elaborated to explore the perception from other disciplines on retractions in the psychology field.

The indicators are statistically determined in each discipline, including the number of CM0, the increment of CM after CM0, as well as the number and the corresponding growth rate of NM before and after CM0. What can be seen from Table 12 is that except psychology itself, *General Social Sciences* have the largest number of CM0, making it the discipline with the highest perception of retractions. After the appearance of CM0, each discipline has various reactions. The NM citations in *Computer Science, Biology & Biochemistry*, and *Pharmacology & Toxicology* notably decrease, whereas in *Psychiatry/Psychology, General Social Sciences, Neuroscience & Behavior*, and *Clinical Medicine*, they still have a slight rise, but simultaneously, CM citations grow as well. The finding suggests that most disciplines, after recognizing the retracted status of references from the psychology field, would have more CM citations and thereby reduce the potential impacts caused by the continued dissemination of retracted papers.

5.3.3. Summary of perception analysis

Theoretically, post-retraction citation is a phenomenon that should not exist, especially the NM citations. As long as the authors, the editors, or the publishers check the references carefully in either part of the entire process from manuscript-writing to article-publishing, the retractions could be found and dealt with. For NM citations, there is a possibility that some authors already knew about the retracted status of the papers, yet they still cited them positively without proper mentioning because the authors approved the methods, standpoint, or findings in the retracted papers. It is hard to entirely figure out the real motivations of these authors' behavior from the empirical data. To alleviate this uncertainty and to understand the scholars' perception of retractions, the unaware duration (UD) of each retracted paper is calculated according to the publication time of the first CM that correctly mentioned the status of retracted papers (CM0), and from which derives two indicators - the growth rate of NM (ΔNM') and the ARC of NM ($k_{\Delta NM0}$, $k_{\Delta NMt}$) - to measure the perception degree. By conducting statistical tests on these indicators, the results are as follows: (1) The overall perception level of the scholars on the retraction in the psychology field is limited, and the unaware duration is nearly 3 years. However, once it has been recognized (i.e. CM0 appears), the rate of NM's increment remarkably slows down. (2) Upon being aware of the retraction

Table 12
CM0 and CM's increment in different disciplines.

No.	Discipline	CM0	CM's increment after CM0	Number of NM before CM0	Number of NM after CM0	ΔNM′
1	Psychiatry/Psychology	18	26	63	93	0.48
2	Social Sciences, General	6	15	19	101	4.32
3	Multidisciplinary	3	3	4	10	1.5
4	Clinical Medicine	3	2	21	28	0.33
5	Neuroscience & Behavior	3	9	39	38	-0.03
6	Economics & Business	2	9	9	24	1.67
7	Environment/Ecology	1	2	0	1	*
8	Plant & Animal Science	1	2	1	0	-1
9	Materials Science	0	0	0	1	\
10	Molecular Biology & Genetics	0	0	2	3	0.5
11	Engineering	0	0	1	3	2
12	Chemistry	0	0	0	1	\
13	Computer Science	0	1	7	1	-0.86
14	Space Science	0	0	1	0	-1
15	Biology & Biochemistry	0	2	5	2	-0.6
16	Physics	0	0	0	1	\
17	Pharmacology & Toxicology	0	0	9	4	-0.56

* \ indicates that the discipline has no NM before CM0, so ΔNM' cannot be calculated.

for the first time, most disciplines consciously mention the retracted status of the reference in the follow-up citations, forming benign dissemination. Hence, it can be further confirmed that the appearance of CM0 has great significance. It means that the unawareness of the scholars has officially come to an end, and the academic community outside the stakeholders of the problematic paper begins to recognize the retraction, taking a series of intentional or unintentional means to break the reference chain.

6. Conclusion

From the different traits of the post-retraction citations of retracted papers in terms of their citation location, citation sentiment, and unaware duration, the scholars' level of perception on retractions in the psychology field is obtained. These traits provide four insights and suggestions for the transmission cycle of retracted papers that are being cited in academia.

First of all, while citing, researchers need to avoid highly controversial literature and also to raise the awareness of updating information by re-downloading papers that have been stored for a long time, as well as confirming the status of references on authoritative and reliable journal websites, or using some technical tools (e.g. CrossMark) to track the correction, erratum and retraction. Furthermore, these services with subscription and alarm functions can update the latest status of papers that readers have downloaded. These methods can help researchers recognize retractions as early as possible. Accordingly, the transmission of retracted papers can be blocked from the beginning, and the cost for later checking and viewing can be reduced.

Secondly, platforms such as journals and publishers are responsible for paper management to further clarify the reasons and motives of post-retraction citations. By doing so, better management can be performed in the tracking and controlling of the spread of retracted papers. On the one hand, the last check before the publication is vital to grasp. With a thorough examination, the problematic references are promptly recognized and eliminated to prevent a resultant malignant spread of the retracted works. On the other hand, papers retracted due to misconduct, unintentional errors, or other reasons have the same tendency of citation sentiment, which is neutral>positive>>negative, causing it difficult to interfere with the continued impacts from the perspectives of retraction reasons. To solve this problem, the citing papers should also explain in-text clearly why they quote the retracted publications, instead of merely marking the retracted status in their references. By doing so, more scholars could recognize it and have chances to make a more reasonable choice, improving the utilization efficiency of scientific research. And for journals, the trustworthiness and accuracy of the scientific studies submitted through an editorial and peer review processes can be ensured (Demir, 2018).

Furthermore, a database with open, dynamically updated, and automatic alerting characteristics built by multi-agents, for example, the Retraction Watch, should be placed on the agenda. Before the authors submit the manuscripts or the journals publish the papers, a comprehensive search and collation in such databases can be helpful in reaching a higher perception level on retractions.

Finally, there is a need for further in-depth interviews with the authors of post-retraction citing papers in order to explore their motivations for not mentioning the retracted status of their references. Fundamentally, it may help to comprehend the contributing factors of the phenomenon and then consummate quoting criteria or paper review guidelines to intervene in the persistent diffusion of retracted works. The constraint of this research, however, is the limited sample scale of 93 retracted papers and 1010 post-retraction citations; the number of retracted papers with CM citations used for final statistical tests is a rather small sample size test of merely 37 papers. It is also worth noting that the unaware duration of different disciplines may be affected by the accumulation and updating speed of each subject (Zhang & Tian, 2019). For instance, in the field of sociology including psychology, the speed of knowledge production is slow, therefore the effectiveness of literature might stay longer. Under this circumstance, the problem of retracted papers spreading through unofficial channels is not easy to be discovered, thus the unaware duration in the field will probably last for a long time. Nevertheless, the same conclusion may not be reached in other disciplines. Hence, the scope of samples needs to be further expanded in the future. It is suggested that the research and comparison should be carried out from multiple dimensions,

such as various disciplines, citation depth, and article types, to enhance the sample data's order of magnitude and to obtain more representative and universal empirical results.

Acknowledgments

This work was financially supported in part by the Ministry of Science and Technology (MOST), Taiwan, under Grant nos.MOST 108-2410-H-002-222-MY3, the Ministry of Education(MOE), Taiwan, under Grant nos. MOE 111L9A002.

References

Teixeira, D., & Bornemann-Cimenti, H. (2017). Why do some retracted papers continue to be cited? *Scientometrics, 110*(1), 365–370. 10.1007/s11192-016-2178-9.
Van, D., & Nijveen, H. (2016). Propagation of errors in citation networks: A study involving the entire citation network of a widely cited paper published in, and later retracted from, the journal Nature. *Research Integrity & Peer Review, 1*(1), 3. 10.1186/s41073-016-0008-5.
Retraction Watch. (2018). Retraction Watch Database User Guide Appendix B: Reasons. Retrieved from https://retractionwatch.com/retraction-watch-database-user-guide/retraction-watch-database-user- guide-appendix-b-reasons/. Accessed July 20, 2020.
Couzin, J. (2006). Scientific misconduct: Cleaning up the paper trail. *Science, 312*(5770), 38–43. 10.1126/science.312.5770.38.
Bar-Ilan, J., & Halevi, G. (2017). Post retraction citations in context: A case study. *Scientometrics, 113*, 547–565. 10.1007/s11192-017-2242-0.
Bornemann-Cimenti, H., Szilagyi, I. S., & Sandner-Kiesling, A. (2016). Perpetuation of retracted publications using the example of the Scott S. Reuben case: Incidences, reasons and possible improvements. *Science and Engineering Ethics, 22*(4), 1063–1072. 10.1007/s11948-015-9680-y.
Neale, A. V., Dailey, R. K., & Abrams, J. (2010). Analysis of citations to biomedical articles affected by scientific misconduct. *Science Engineering Ethics, 16*(2), 251–261. 10.1007/s11948-009-9151-4.
Furman, J. L., Jensen, K., & Murray, F. (2012). Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. *Research Policy, 41*(2), 276–290. 10.1016/j.respol.2011.11.001.
Chen, C. M., Hu, Z. G., Milbank, J., et al., (2012). A visual analytic study of retracted articles in scientific literature. *Journal of American Society for Information Science and Technology, 64*(2), 234–253. 10.1002/asi.22755.
Pfeifer, M. P., & Snodgrass, G. L. (1990). The continued use of retracted, invalid scientific literature. *JAMA, 263*(10), 1420–1423.
Hamilton, D. G. (2019). Continued citation of retracted radiation oncology literature—Do we have a problem? *International Journal of Radiation Oncology*Biology*Physics, 103*(5), 1036–1042. 10.1016/j.ijrobp.2018.11.014.
Budd, J. M., Sievert, M., & Schultz, T. R. (1998). Phenomena of retraction: Reasons for retraction and citations to the publications. *JAMA, 280*, 296–297. 10.1001/jama.280.3.296.
Kochan, C. A., & Budd, J. M. (2010). The persistence of fraud in the literature: The Darsee case. *Journal of the Association for Information Science & Technology, 43*(7), 488–493 https://doi.org/10.1002/(SICI)1097-4571(199208)43:7<488::AID-ASI3>3.0.CO;2-7.
Neale, A. V., Northrup, J., Dailey, R., et al., (2007). Correction and use of biomedical literature affected by scientific misconduct. *Science and Engineering Ethics, 13*(1), 5–24. 10.1007/s11948-006-0003-1.
Nath, S. B., Marcus, S. C., & Druss, B. G. (2006). Research enterprise retractions in the research literature: Misconduct or mistakes? *Medical Journal of Australia, 185*(3), 152. 10.5694/j.1326-5377.2006.tb00504.x.
Wager, E., Barbour, V., Yentis, S., & Kleinert, S. (2009). Retractions: guidance from the Committee on Publication Ethics (COPE). *Croatian medical journal, 50*(6), 532–535. 10.3325/cmj.2009.50.532.
Tian, R. Q, Yao, C. Q., Liu, Y. Y., et al., (2014). Research on the conditions, reasons and procedures for the retraction of academic papers. *Chinese Journal of Scientific and Technical Periodicals, 25*(10), 1232–1237.
Wright, K., & Mcdaid, C. (2011). Reporting of article retractions in bibliographic databases and online journals. *Journal of the Medical Library Association (JMLA), 99*(2), 164–167. 10.3163/1536-5050.99.2.010.
Suelzer, E. M., Deal, J., Hanus, K. L., et al., (2019). Assessment of citations of the retracted article by wakefield et al. with fraudulent claims of an association between vaccination and autism. *JAMA Network Open, 2*(11), Article e1915552. 10.1001/jamanetworkopen.2019.15552.
Rgac, D., Ag, A., & Mka, B. (2019). Why articles continue to be cited after they have been retracted: An audit of retraction notices-ScienceDirect. *International Journal of Nursing Studies, 90*, 11–12. 10.1053/j.jvca.2021.09.005.
Teixeira, D., & Dobránszki, J. (2017). Highly cited retracted papers. *Scientometrics, 110*, 1653–1661. 10.1007/s11192-016-2227-4.
Kim, S. Y., Yi, H. J., Cho, H. M., et al., (2019). How many retracted articles indexed in KoreaMed were cited 1 year after retraction notification. *Science Editing, 6*(2), 122–127. 10.6087/kcse.172.
Budd, J. M., Sievert, M., Schultz, T. R., et al., (1999). Effects of article retraction on citation and practice in medicine. *Bulletin of the Medical Library Association, 87*(4), 437–443.
Moravcsik, M. J., & Murugesan, P. (1975). Some results on the function and quality of citations. *Social Studies of Science, 5*(1), 86–92. 10.1177/030631277500500106.
Chubin, D. E., & Moitra, S. D. (1975). Content analysis of references: Adjunct or alternative to citation counting? *Social Studies of Science, 5*(4), 423–441. 10.1177/030631277500500403.
Liu, S. B., Chen, C. M., Ding, K., et al., (2014). Literature retrieval based on citation context. *Scientometrics, 101*(2), 1293–1307. 10.1007/s11192-014-1233-7.
Zhang, G., Ding, Y., & Milojevi, S. (2013). Citation content analysis (CCA): A framework for syntactic and semantic analysis of citation content. *Journal of the American Society for Information Science and Technology, 64*(7), 1490–1503. 10.1002/asi.22850.
Wright, N. D. (1991). *A citation context analysis of retracted scientific articles*. College Park: University of Maryland.
Korpela, K. M. (2010). How long does it take for the scientific literature to purge itself of fraudulent material?: the Breuning case revisited. *Current Medical Research and Opinion, 26*(4), 843–847. 10.1185/03007991003603804.
Catalini, C., Lacetera, N., & Oettl, A. (2015). The incidence and role of negative citations in science. *Proceedings of the National Academy of Sciences of the United States of America, 112*(45), 13823–13826. 10.1073/pnas.1502280112.
Zhao, R. Y., Zeng, X. Q., & Chen, B. K. (2014). Citation in full-text: The development of citation analysis. *Library and Information Service, 58*(9), 129–135 http://doi.org/10.13266/j.issn.0252-3116.2014.09.018.
Luong, M. T., Nguyen, T. D., & Kan, M. Y. (2010). Logical structure recovery in scholarly articles with rich document features. *International Journal of Digital Library Systems, 1*(4), 1–23. 10.4018/jdls.2010100101.
An, J., Kim, N., Kan, M. Y., et al., (2017). Exploring characteristics of highly cited authors according to citation location and content. *Journal of the Association for Information Science and Technology, 68*(8), 1975–1988. 10.1002/asi.23834.
Pak, C. M., Wang, W., & Yu, G. (2020). An analysis of in-text citations based on fractional counting. *Journal of Informetrics, 14*(4), Article 101070. 10.1016/j.joi.2020.101070.
Zhang, C., Liu, L., & Wang, Y. (2021). Characterizing references from different disciplines: A perspective of citation content analysis. *Journal of Informetrics, 15*(2), Article 101134. 10.1016/j.joi.2021.101134.
Teufel, S. (1999). *Argumentative zoning: information extraction from scientific text.* Edinburgh: University of Edinburgh.
Weinstein, N. D., Sandman, P. M., & Blalock, S. J. (2008). The precaution adoption process model. *Health Psychology, 7*(4), 123–147. 10.1002/9781119057840.ch100.
Demir, S. B. (2018). Predatory journals: Who publishes in them and why? *Journal of Informetrics, 12*(4), 1296–1311. 10.1016/j.joi.2018.10.008.
Zhang, L. H., & Tian, D. (2019). Comparative analysis of the retraction reasons and time-to-retraction in different disciplines. *Digital Library Forum, 2019*(8), 40–44.