Postprint: Improving the Disruptiveness Index for Research Papers and Its Influencing Factors

Authors: Liu Xiaohui, Shen Zhesi, Liao Yu, YANG Liying

Date: 2023-04-01T16:16:05+00:00

Abstract

[Purpose/Significance] To resolve the inconsistency issues in the mathematical properties of the disruptive index and investigate the influencing factors in its application. [Method/Process] First, the inconsistent behaviors of the disruptive index D are identified; then, improved versions are developed as the relative disruptive index Rela_{DZ} and the absolute disruptive index DZ; finally, factors influencing differences in disruptive indices are analyzed from three perspectives: citation time window, disciplinary differences, and document type. [Results/Conclusion] The Rela_{DZ} algorithm resolves the non-monotonicity problem of D with respect to NR, while the DZ algorithm resolves the not-strictly-monotonic problem of D with respect to NF and NB, thereby eliminating inconsistency; combining the relative and absolute algorithms for the disruptive index proves more reasonable in application. Furthermore, the disruptive indices Rela_{DZ} and DZ are influenced by citation time window, discipline, and document type, therefore, appropriate adjustments should be made in their application.

Full Text

Preamble

Title: Research on the Improvement of the Disruption Index for Scientific Papers and Its Influencing Factors

Authors: Liu Xiaohui[1,2], Shen Zhesi[1], Liao Yu[1,2], Yang Liying[1,2]

Affiliations:
[1] National Science Library, Chinese Academy of Sciences, Beijing 100190, China
[2] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China

Abstract:
[Purpose/Significance] This study addresses the inconsistency problem in the mathematical properties of the disruption index and explores the factors influencing its application. [Method/Process] We first identify the inconsistency issues in the original disruption index D, then propose improved versions: the relative disruption index Rela_{DZ} and the absolute disruption index DZ. Finally, we analyze the factors affecting these indices from three perspectives: citation time window, disciplinary differences, and document types. [Result/Conclusion] The Rela_{DZ} algorithm resolves the non-monotonicity of D with respect to NR, while the DZ algorithm solves the non-strict monotonicity issues of D with respect to NF and NB, thereby eliminating inconsistency. Combining both relative and absolute algorithms yields more reasonable results in practice. Additionally, both Rela_{DZ} and DZ are influenced by citation time window, discipline, and document type, necessitating appropriate adjustments in application.

Keywords: disruption index; creativity; inconsistency; scientific evaluation; citation analysis
Classification Number: G250.2
DOI: 10.13266/j.issn.0252-3116.2020.24.010

As scientific research continues to develop, evaluation methods have evolved toward greater fairness and rationality. The core of this evolution has shifted from focusing solely on quantity to balancing quality and quantity, and from emphasizing impact to prioritizing innovation. The novelty and innovativeness of research papers have gradually gained recognition and become a hotspot in evaluation studies.

Current research on innovation measurement can be categorized into three main approaches. The first defines innovativeness or novelty from the perspective of knowledge combination. For instance, B. Uzzi et al. [2], Y. Lee et al. [3], J. Wang et al. [4], and Y. He et al. [5] measure novelty through atypical combinations of journals in reference lists; K. Boyack et al. [6] and R. Klavans et al. [7] construct disciplinary maps based on disciplinary combinations and measure innovation through inter-disciplinary distances; N. Carayol et al. [8], Wang Yanyan et al. [9], Lu Wanhui et al. [10], Shen Yang [11], and Yang Jianlin et al. [12] assess paper novelty from co-word and textual semantic perspectives using keywords and content analysis. The second approach defines novelty temporally. For example, S. Mishra et al. [13], M. Packalen et al. [14], and E. Callaway [15] measure article novelty through concept age—the time elapsed since concepts mentioned in the paper were first introduced—where later concept introduction indicates greater novelty. The third approach uses citation structure to measure innovation, such as D. Trapido [16], R. Funk et al. [17], and L. Wu et al. [18], who measure the degree to which papers replace prior knowledge and influence subsequent research through local citation structures, thereby defining disruptiveness.

The disruption index, as an important metric for measuring innovation, has attracted widespread attention in the scientometrics community since its proposal, with scholars also noting its inherent inconsistency problems. This paper proposes improved algorithms for the disruption index and systematically analyzes the performance and influencing factors of the improved indices.

2. Related Research on the Disruption Index

The most widely discussed innovation metric in current domestic and international research is the disruption index D (Disruption Index) [18], which has been extensively explored by scholars. Some studies focus on its application. For example, L. Bornmann et al. [19] applied the D algorithm to calculate disruption indices for papers published in *Scientometrics* from 2000 to 2010, finding that most indices clustered around 0 with only a few showing significant variation. Bornmann et al. also explored the computational requirements of index D, analyzing four papers from different disciplines and discovering that the disruption index depends on the citation time window [20], requiring at least three years to stabilize. Other studies have examined the D algorithm itself. For instance, L. Bornmann et al. questioned whether the original disruption index algorithm truly measures disruptiveness, proposed a family of new disruption index algorithms DI_n [21], and applied DI_5 to recalculate disruption indices for *Scientometrics* papers [22], finding that DI_5 better identifies disruptive research that captures the attention of researchers within disciplinary fields. Q. Wu et al. utilized elements from the original disruption index algorithm to propose four similar indicators, identifying three relatively reasonable algorithms for calculating disruption indices [23].

However, existing research largely neglects the properties of the disruption index or its influencing factors in application. Although some scholars have explored the stabilization time window for the index, their analysis was based on only four papers, which lacks statistical generality and representative conclusions.

2.1 Conceptual Introduction to the Disruption Index

The disruption index D was originally proposed by R. Funk et al. [17] for patents and later extended to papers and computer programs by L. Wu et al. [18] in 2019. Their work revealed that small teams produce more disruptive research while large teams tend to produce developmental research, a finding featured on the cover of *Nature*. The calculation method is shown in Equation (1):

$$D = p_F - p_B = \frac{N_F - N_B}{N_F + N_B + N_R}$$

In Equation (1), all statistics are measured after the publication of the focal paper (the paper being evaluated). L. Wu et al. categorized subsequent research into three types (see Figure 1 [Figure 1: see original paper]): Type F cites only the focal paper; Type B cites both the focal paper and its references; Type R cites only the focal paper's references but not the focal paper itself. Based on this classification, the disruption index is defined as the difference in proportions between Type F and Type B studies. In the formula, N_F, N_B, and N_R represent the numbers of Type F, Type B, and Type R citing papers, respectively.

The value range of D is $[-1, 1]$. $D > 0$ indicates the focal paper is disruptive; $D < 0$ indicates it is developmental. When $N_B = N_R = 0$, $D = 1$, representing complete disruption of prior research. When $N_F = N_R = 0$, $D = -1$, representing complete development or consolidation of prior research. When $N_F = N_B$, $D = 0$, indicating neutrality.

2.2 Inconsistency of the Disruption Index

Although widely recognized, research has revealed that the disruption index exhibits inconsistency in its mathematical properties, potentially leading to inaccurate conclusions in practical applications.

"Inconsistency" is a concept opposite to "cognitive consistency theory" [24] in social psychology. In scientometrics, inconsistency can be understood as: when an independent variable changes in an indicator, the dependent variable changes in a way that contradicts intuitive expectations. Scholars have studied inconsistencies in various metrics, such as the A-index [25], H-index [26], and journal impact factor [23]. According to L. Egghe et al. [27], inconsistency includes violations of monotonicity and relative independence.

The inconsistency of the disruption index D refers to its non-monotonic behavior with respect to N_R. S. Wu et al. first pointed out this inconsistency at the virtual case level [28] but did not systematically reveal its causes or propose solutions.

Specifically, D's non-monotonicity with respect to N_R means: when $N_F > N_B$, D decreases monotonically with N_R; when $N_F < N_B$, D increases monotonically with N_R; but when $N_F = N_B$, D becomes non-monotonic with respect to N_R, as shown in Table 1 . In the disruption index formula, N_R represents "the number of citing papers after the focal paper's publication that cite only the focal paper's references, not the focal paper itself." Such citations inherit or develop the references, reflecting their importance rather than the focal paper's. Intuitively, larger N_R should indicate lower disruptiveness of the focal paper. However, when $N_F < N_B$, D increases with N_R, creating inconsistency.

Table 1 illustrates this inconsistency through examples. Mathematically, the derivative of D with respect to N_R is:

$$\frac{\partial D}{\partial N_R} = \frac{N_B - N_F}{(N_F + N_B + N_R)^2}$$

As shown in Equation (2), the sign of D's derivative with respect to N_R depends on the relative magnitude of N_B and N_F. When $N_F > N_B$, the derivative is negative; when $N_F < N_B$, the derivative is positive.

3. Improved Algorithms for the Disruption Index

3.1 Defining Relevant Factors for Improved Disruption Index Algorithms

L. Wu et al. categorized citations into three types: F, B, and R [18]. Our improvements adopt this substitution-based framework, making these three citation types crucial factors. We define their meanings as follows: for characterizing the focal paper, Type F citations reflect its disruptiveness to references, Type B indicates its development of references, and Type R reveals its inheritance from references. Thus, F, B, and R citations represent the disruptiveness, developmental nature, and inheritance of the focal paper, respectively.

Since the original disruption index D measures disruptiveness based on the relative magnitude of N_F and N_B, we consider it a "relative" concept. Correspondingly, we define the disruptiveness measured solely by Type F citations as the "absolute disruption index."

3.2 Absolute Disruption Index DZ

From an "absolute" perspective and based on the meanings of the three citation types, we improve the disruption index D through three considerations: First, a paper's disruptiveness is manifested only in Type F citations. Second, as the number of Type F citations increases, the 偶然性 (randomness) of generating them decreases while the 必然性 (inevitability) increases, meaning the growth rate of disruptiveness should accelerate with more Type F citations rather than remain constant. Third, since Type R citations may differ significantly from the focal paper's research topic, their impact on the disruption index should be appropriately smaller than that of the other two types.

Based on these considerations, we derive the improved algorithm DZ:

$$DZ = \frac{2N_F}{2N_F + 2N_B + N_R}, \quad DZ \in [0, N_F]$$

DZ measures disruptiveness using only Type F citations, capturing the impact of the focal paper's disruptive characteristics on scientific development. In contrast, algorithm D measures relative disruptiveness by comparing Type F and Type B citations. Therefore, we designate DZ as the absolute disruption index and D as the relative disruption index. If $N_F = 0$, then DZ = 0 regardless of other variables, while D may be negative.

3.3 Determining Relative Disruptiveness Tendency

Since DZ measures only absolute disruptiveness, the presence of "neutral" citing papers (Type R) also reduces the index value. Thus, we define the opposite of "tending toward disruption" as "not tending toward disruption"—including both "tending toward development" and "neutral."

Assuming $N_R = 0$ and denoting the total citations as C:

1. When $N_F \leq N_B$, the paper does not tend toward disruption: $DZ \leq \frac{2N_F}{2N_F+2N_B} = \frac{N_F}{N_F+N_B}$, where $C = N_F + N_B$
2. When $N_F > N_B$, the paper tends toward disruption: $DZ > \frac{2N_F}{2N_F+2N_B} = \frac{N_F}{N_F+N_B}$, where $C = N_F + N_B$
3. Since Type R citations reflect inheritance, their presence further lowers the threshold for judging relative disruption. Therefore, the threshold when $N_R = 0$ represents the maximum boundary for general cases and remains applicable when $N_R \neq 0$.

Combining (1), (2), and (3), $C/4$ serves as the threshold for determining relative disruptiveness tendency. When $DZ = C/4$, the focal paper is neutral.

3.4 Relative Disruption Index Rela_{DZ}

Under the DZ design framework, to measure the degree of relative disruptiveness tendency, we express relative disruptiveness as DZ's relative value to this threshold, denoted as Rela_{DZ}. To eliminate differences caused by citation frequency, the algorithm is:

$$Rela\_DZ = \frac{4DZ}{C} = \frac{8N_F}{2C^2 + C \cdot N_R}, \quad Rela\_DZ \in [0, 4]$$

Table 2 shows the disruption indices for the papers in Table 1 under DZ and Rela_{DZ}:

Table 2. Validation of Improved Indicators

NF and NB	NF > NB	NF = NB
NF NB NR	DZ Rela_{DZ}	About NR
90 10 0	81.00 3.24	Monotonically decreasing
90 10 100	54.00 2.16	Monotonically decreasing
10 90 0	1.00 0.04	Monotonically decreasing
10 90 100	0.67 0.03	Monotonically decreasing
10 10 0	5.00 1.00	Monotonically decreasing
10 10 100	1.43 0.29	Monotonically decreasing

As shown in Table 2, whether $N_F > N_B$ or not, both DZ and Rela_{DZ} decrease monotonically with N_R, satisfying consistency.

3.5 Validation of Improvement Effect

We validate the improvement effect using papers from the American Physical Society (APS) journals and selected milestone papers. The APS's *Physical Review* series represents important journals in the physics community, with *Physical Review Letters* (PRL) recognized as a top-tier journal publishing frontier work with significant impact. In 2008, the APS editorial board selected a collection of APS milestone papers for the society's 50th anniversary. Our dataset includes 548,133 APS papers, 115,648 PRL papers, and 75 milestone papers.

Table 3 shows the mean percentile rankings of disruption indices across algorithms. Given the dataset characteristics, average paper quality increases from APS to PRL to milestone papers, so their disruption rankings should correspondingly improve. Under Rela_{DZ} and DZ, all three datasets show this expected pattern, whereas under D, PRL papers' rankings unexpectedly decrease. For milestone papers, rankings under DZ and Rela_{DZ} show substantial improvement compared to D.

Unlike the original index D, the improvement process provides DZ from an absolute perspective, which can be used independently to examine the impact and 推动作用 (driving effect) of a focal paper's disruptive characteristics on scientific development. This differs from R. Funk's simple product with citation frequency [17].

In summary, Rela_{DZ} improves upon D's non-monotonicity with respect to N_R and, like D, belongs to relative metrics suitable for assessing a paper's disruptive tendency without considering scale. DZ offers a new perspective, reflecting the impact scale of a paper's disruptive attributes in the scientific community, applicable independently. DZ is monotonic with respect to N_R and strictly monotonic with respect to N_F and N_B, making it an absolute metric. Combined use of Rela_{DZ} and DZ provides more comprehensive and reasonable revelation of a paper's disruption level.

4. Influencing Factors of DZ and Rela_{DZ}

Citation-based metrics are typically affected by time windows and document types. Additionally, disciplinary differences in knowledge characteristics, research nature, and citation behavior cause variations across fields. This section explores these three factors' effects on DZ and Rela_{DZ} to provide recommendations for proper application.

4.1 Citation Time Window

The disruption index is dynamic, depending on citations. Since citation accumulation requires time, a stable citation count is needed to approximate a paper's true disruption level. Different disciplines require different time windows for citations to stabilize. We employ dynamic citation windows to examine their impact on DZ and Rela_{DZ}, using citation frequency trends as a comparative baseline.

Using 2007 SCI papers from Web of Science (with references dating back to 1900), we calculated disruption indices for citation windows of 2, 4, 6, 8, and

10 years. Papers from computer science, engineering, economics & business, and social sciences were excluded due to their predominant coverage in CPCi or SSCI, yielding 795,022 focal papers.

Figures 2 [Figure 2: see original paper] and 3 [Figure 3: see original paper] show DZ and Rela_{DZ} trends across disciplines, while Figure 4 [Figure 4: see original paper] displays citation frequency trends (all using mean values).

Analysis reveals: (1) After four years post-publication, citation frequencies across disciplines diverge significantly, becoming more pronounced after six years. These differences in citation frequency create substantial gaps in both DZ and Rela_{DZ} across time windows. Thus, citation windows affect disruption indices through citation frequency, requiring sufficient time for stabilization. (2) The growth rate of citation frequencies is lower than that of both disruption indices, indicating that temporal accumulation creates differential growth among the three citation types (F, B, R), directly impacting disruption indices more strongly than citation counts alone.

4.2 Disciplinary Differences in Disruption Indices

Citation metrics generally exhibit disciplinary differences. While DZ's numerator and denominator have different dimensions (unlike Rela_{DZ}), both show distribution variations. Using 2015 SCI papers (1,318,359 focal papers) with a 2-year citation window, we explored these differences through cumulative distribution plots and statistical tests.

Cumulative distribution functions (Figure 5 [Figure 5: see original paper]) reveal clear disciplinary groupings for DZ: (1) molecular biology & genetics, space science, neuroscience, microbiology, biology & biochemistry, chemistry; (2) psychiatry, physics, plant & animal science, geosciences, environmental ecology, pharmacology & toxicology, immunology; (3) clinical medicine, materials science, agriculture; and (4) mathematics. For Rela_{DZ}, differences are smaller overall, though the proportion of papers at the minimum DZ value varies substantially across disciplines (75% to nearly 85%).

To quantify these differences, we applied non-parametric tests since the data violate normality (K-S test, $p < 0.01$). Table 4 shows the proportion of disciplinary pairs with significant differences ($p < 0.001$) using both Kolmogorov-Smirnov (K-S) and Kruskal-Wallis (K-W) tests. Both methods confirm significant disciplinary differences for DZ and Rela_{DZ} at the 0.001 level.

4.3 Document Types

Article and Review are the most common document types. Articles typically represent original research with new discoveries, while Reviews synthesize and summarize prior work, generally lacking original scientific knowledge. Intuitively, Articles should have higher disruptiveness.

Using 2015 SCI papers (1,241,475 Articles and 76,884 Reviews), we examined this effect. Table 5 presents means and 95% confidence intervals (via bootstrapping) for both indices.

Table 5. Mean Disruption Indices for Review and Article Papers

Document Type	DZ (95% CI)	Rela_{DZ} (95% CI)
Reviews	0.983 (0.76, 1.28)	0.054 (0.047, 0.057)
Articles	0.541 (0.49, 0.60)	0.137 (0.132, 0.139)

Results show: (1) For absolute disruptiveness (DZ), Reviews have higher values, partly due to non-normal distributions and smaller sample sizes, but primarily because DZ correlates positively with citation frequency (Reviews: 12.23 citations; Articles: 5.66 citations). This difference reflects citation impact rather than greater scientific innovativeness, making cross-type comparisons inappropriate. (2) For relative disruptiveness (Rela_{DZ}), Articles show higher values, reflecting their essential nature: Articles generate relatively more Type F citations, indicating greater originality. (3) Document type significantly affects both absolute and relative disruption indices.

5. Discussion and Conclusion

The disruption index measures a focal paper's replacement of prior knowledge from the perspective of citing literature to assess its originality. L. Wu et al.'s algorithm D expresses a relative concept with clever design but suffers from mathematical inconsistency that may produce misleading conclusions. Addressing this, we propose Rela_{DZ}, which resolves D's non-monotonicity with respect to N_R while maintaining the relative perspective. The improvement process also yields DZ, an absolute metric reflecting the impact scale of a paper's disruptive attributes, which can be used independently. DZ is monotonic with respect to N_R and strictly monotonic with respect to N_F and N_B.

Every metric has influencing factors and application characteristics. To enable comprehensive understanding and proper application, we examined three common aspects of citation metrics: citation time window, disciplinary differences, and document types. All three factors affect the indices, fundamentally due to their non-independent relationship with citation frequency—a topic requiring further detailed investigation.

The improved disruption indices share limitations of citation-based metrics: they cannot distinguish citation attitudes, are affected by time windows, and cannot evaluate zero-cited papers. However, they also offer advantages like resistance to manipulation. Overall, the disruption index represents progress in research evaluation methodology. Methodologically, this study provides a framework for addressing inconsistency issues in scientometric indicators and offers a more scientifically sound quantitative metric for research evaluation, helping reveal scientific activity patterns and supporting fairer, more efficient research management.

References:

[1] Li Zhimin. Reform of China's S&T evaluation: How has it evolved? What are the trends? [EB/OL]. [2019-08-29]. https://mp.weixin.qq.com/s/if_{1621}.002.html.
[2] Uzzi B, Mukherjee S. Atypical combinations and scientific impact [J]. Research policy, 2015, 44(3): 684-697.
[3] Lee Y, Walsh JP, Wang J. Creativity in scientific teams: unpacking novelty and impact [J]. Science, 2013, 342(6157): 468-472.
[4] Wang J, Veugelers R, Stephan P. Bias against novelty in science: a cautionary tale for users of bibliometric indicators [J]. D-Lib magazine, 2016, 22: 9-10.
[5] He Y, Luo J. Novelty, conventionality, and value of invention [M]//Design computing and cognition'16. Cham: Springer International Publishing, 2017: 23-38.
[6] Boyack K, Klavans R. Is the most innovative research being funded? [EB/OL]. [2020-12-08]. https://www.elsevier.com/___{data}/assets/pdf_{file}/0005/479426/boyack_{{{sti}}
[7] Klavans R, Boyack K. Towards the development of an article-level indicator of conformity, innovation and deviation [C]//Proceedings of 18th international conference on science and technology indicators. Berlin: STI, 2013: 185-192.
[8] Carayol N, Lahatte A, Llopis O. Novelty in science presented at STI2017 [EB/OL]. [2020-12-08]. https://digital.csic.es/bitstream/10261/162613/1/novelty.pdf.
[9] Wang Yanyan, Zhang Junsheng, Qiao Xiaodong, et al. A literature novelty assessment method based on problem-method matrix [J/OL]. Information Studies: Theory & Application: 1-13. [2020-11-03]. http://kns.cnki.net/kcms/detail/11.1762.G3.20201030.1621.002.htm
[10] Lu Wanhui, Tan Zongying. Research on measuring topic novelty of academic achievements based on Doc2Vec and HMM algorithm [J]. Data Analysis and Knowledge Discovery, 2018, 2(3): 22-29.
[11] Shen Yang. A keyword-based method for evaluating innovation degree [J]. Information Studies: Theory & Application, 2007(1): 125-127.
[12] Yang Jianlin, Qian Lingfei. Measuring topic novelty based on inverse document frequency of keyword pairs [J]. Information Studies: Theory & Application, 2013, 36(3): 99-102.
[13] Mishra S, Torvik V. Quantifying conceptual novelty in the biomedical literature [J]. Research policy, 2015, 44(3): 684-697.
[14] Packalen M, Bhattacharya J. Citations and ideas [EB/OL]. [2020-10-16]. http://www.nber.org/papers/w20921.
[15] Callaway E. Young scientists lead the way on fresh ideas [J]. Nature, 2015, 7539(518): 283.
[16] Trapido D. How novelty in knowledge earns recognition: the role of consistent identities [J]. Research policy, 2015, 8(44): 1488-1500.
[17] Funk R, Smith J. A dynamic network measure of technological change [J]. Management science, 2017, 63(3): 791-817.
[18] Wu L, Wang D, Evans J. Large teams develop and small teams disrupt science and technology [J]. Nature, 2019, 566(7744): 378-382.

[19] Bornmann L, Tekles A. Disruptive papers published in Scientometrics [J]. Scientometrics, 2019, 120(1): 331-336.

[20] Bornmann L, Tekles A. Disruption index depends on length of citation window [J]. El profesional de la información, 2019, 28(2): 1-2.

[21] Bornmann L, Devarakonda S, Tekles A, et al. Do disruption index indicators measure what they propose to measure? The comparison of several indicator variants with assessments by peers [EB/OL]. [2020-03-20]. https://arxiv.org/abs/1911.08775.

[22] Bornmann L, Devarakonda S, Tekles A, et al. Disruptive papers published in Scientometrics: meaningful results by using an improved variant of the disruption index originally proposed by Wu, Wang, and Evans (2019) [J]. Scientometrics, 2020, 123(2): 1149-1155.

[23] Wu Q, Yan Z. Solo citations, duet citations, and prelude citations: new measures of the disruption of academic papers [EB/OL]. [2020-02-21]. https://www.researchgate.net/publication/332977541{{{Solo}}*{{citations}}*}{{{duet}}*{{citations}}*}{{{and}}

[24] Cognitive consistency theory [EB/OL]. [2020-02-21]. https://wiki.mbalib.com/wiki/认知相符理论.

[25] Rousseau R. The F-measure for research priority [J]. Journal of data and information science, 2018(3): 1-18.

[26] Waltman L, Van Eck N. The inconsistency of the h-index [J]. Journal of the American Society for Information Science and Technology, 2012, 63(2): 406-415.

[27] Egghe L, Rousseau R. A general framework for relative impact indicators [J]. Canadian journal of information and library science-revue canadienne des sciences de l'information et de bibliothéconomie, 2002, 27(1): 29-48.

[28] Wu S, Wu Q. A confusing definition of disruption [EB/OL]. [2020-12-29]. https://osf.io/preprints/socarxiv/d3wpk/.

Author Contributions:

Liu Xiaohui: Designed the paper structure, collected and processed data, wrote the initial draft.

Shen Zhesi: Supervised paper structure, revised the manuscript.

Liao Yu: Processed data, created visualizations, revised the manuscript.

Yang Liying: Supervised paper structure, revised the manuscript.

Note: Figure translations are in progress. See original paper for figures.