Information Processing and Management 61 (2024) 103718

Contents lists available at ScienceDirect

Information Processing and Management

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

Decoding the writing styles of disciplines: A large-scale quantitative analysis

Shuyi Dong [a], Jin Mao [b,*], Qing Ke [c], Lei Pei [a]

[a] School of Information Management, Nanjing University, Nanjing, PR China
[b] Center for Studies of Information Resources, Wuhan University, Wuhan, PR China
[c] School of Data Science, City University of Hong Kong, Hong Kong, PR China

A R T I C L E I N F O

Keywords:
Writing style
Interpretable machine learning
Scientific culture
Linguistic features
Shap

A B S T R A C T

Disciplinary writing style stems from the practice of science, reflecting the scientific culture. This study aims to explore the differences and evolution of scientific writing styles from the perspective of disciplines. A large-scale quantitative analysis was conducted over 14 million abstracts from the Microsoft Academic Graph (MAG) database across eight soft and hard disciplines. Represented by a comprehensive set of 14 symbolic, lexical, syntactic, structural, and readability features, the evolution of disciplinary writing styles was analyzed over 30 years. Interpretable machine learning methods were performed to test the discernibility of writing styles across disciplines and disclose their linguistic differences. Our findings reveal the linguistic features of soft disciplines (Art, Philosophy, and Sociology) and Mathematics generally keep stabilized, and a general trend of increasing linguistic complexity was observed for Biology, Chemistry, Computer Science, and Psychology. The good performance of the pairwise writing style classifiers indicates a well discriminability of the writing styles between disciplines. A correlation between the performance of classifiers and the distance between disciplines was identified. The feature contribution analysis using SHapley Additive exPlanations (SHAP) and Kendall's Tau rank correlation revealed the detailed commonalities and disparities in disciplines' linguistic features. This study provides profound insights into the understanding of scientific writing and norms, which further helps develop useful tools for academic text analysis, foster interdisciplinary communication, and assist educators to construct discipline-specific writing guidance.

1. Introduction

"*Academics do not act in a social vacuum and knowledge is not constructed outside particular communities of practice.*" (Hyland, 2006, p. 39) As an essential transmitter of scientific findings, scientific papers convey information, establish scholarly communities of shared knowledge (Harrison & Stephen, 1995), and lay the foundation for scientific communication and development (Abelson, 1980). In addition to the content of scientific writing, its writing style is also of great significance to scientific communication, as authors will choose appropriate discipline-sensitive linguistic resources and present their studies to intended readers and reviewers (Hyland, 2006).

The style of scientific writing can be seen as a reflection of science culture. The original culture of science is characterized by

* Corresponding author.
E-mail address: danveno@163.com (J. Mao).

https://doi.org/10.1016/j.ipm.2024.103718
Received 27 August 2023; Received in revised form 15 February 2024; Accepted 10 March 2024
Available online 23 March 2024

independence of thought and dissent (Iaccarino, 2003), while external societal and institutional factors shape its practices (Smaldino et al., 2019). At the discipline level, culture consists of knowledge traditions in which scholars from the same field share updated findings, hold similar cultural notions, and follow uniform codes of conduct (Berkenkotter et al., 1991; Harrison & Stephen, 1995). Under disciplines' consistent research goals, genres, and vocabularies, unified sets of communication symbols were formed, manifesting as similarities in writing (Swales, 1990). Humanities and social sciences, for instance, prioritize analysis and synthesis of diverse sources (Winch, 2015), whereas natural sciences and engineering disciplines emphasize experimental content such as defining objects, describing procedures, and reporting results (Sletten & Bertozzi, 2009). Educators have made substantial attempts to assist novices in acquiring specialized literacies pertinent to their disciplines (Zhu, 2004). Yet, for researchers, academic writing is a skill developed through practice and intuition, informed by mentors' guidance and peer reviews' feedback. While general guidelines dictated how to write academically, novice researchers might still need to spend considerable time learning and mastering their discipline's specific writing nuances for successful publication.

In the age of increasing interdisciplinary collaboration, many complex problems need to be solved through collaboration among distinct disciplines (Larivière et al., 2015). However, their varied writing styles might hamper their cooperation (Demarest & Sugimoto, 2015; Gonsalves et al., 2021). When scholars from various fields contribute to a single paper, the disparate styles can lead to issues of coherence and flow, necessitate additional adjustments, and potentially cause disagreements or insufficient recognition of others' contributions. Studies have shown that the more divergent the disciplines are, the greater the likelihood that their collaboration will yield successful and influential outcomes (Choi & Pak, 2008; Larivière et al., 2015). However, when disciplines, knowledge systems, and cultural backgrounds vary substantially, it is doubtful that the collaborators will share the same writing criteria. Consequently, they may face the challenge of appreciating an alien scientific culture and speaking foreign scientific languages, which is the so-called "*translation problem*" (Salter & Hearn, 1997, p.59).

Nevertheless, this problem has not yet been sufficiently investigated. Influenced by universalism (Merton, 1974), scientific writing is often considered as a unified concept in contrast with general writing. Relevant studies predominantly focus on the resemblances and correspondence between scientific texts from different disciplines, rather than exploring their differences (Hyland, 2006). However, the differences between the styles of disciplines can be perceived by readers as differences in readability, abstractness, informativeness, etc. (Kaslow, 2015; Qiu et al., 2022). Therefore, it is essential to deepen our understanding of the *translation problem*, identify the differences in writing styles among disciplines, disclose the underlying driving forces, and thereby facilitate interdisciplinary communication and advance scientific progress.

Several scientific discourse studies using mixed research methods, including quantitative text analysis, genre analysis (Matthews & Glitre, 2021), and rhetorical analysis (Li & Jiao, 2022),have provided convincing results in understanding disciplinary writing. Quantitative studies have explored variations in writing styles through register theory (Halliday, 1978) and epistemology (Demarest & Sugimoto, 2015; Li & Jiao, 2022). Some have identified significant measurable differences in linguistic traits across disciplines using machine learning and statistical methods, finding disciplinary differences directly linked to methodological variations (Argamon et al., 2008; Teich et al., 2016). However, these studies generally fell short of depicting a panoramic pattern of disciplines' writing styles. This leaves significant gaps in our understanding of the differences in disciplines' writing styles and how they relate to the broader cultural contexts of each discipline in a broader landscape.

Furthermore, the increasing use of AI tools (e.g., ChatGPT[1]) in refining scientific manuscripts raises pivotal questions about the blurring boundaries and evolving paradigms in academic authorship. Therefore, it is crucial to establish a foundational framework to reveal and capture the characteristics of academic discourses across disciplines. Such a framework could enhance the understanding of the consistency that shaped by the long-term adherence to specific norms and regulations within each discipline.

To address these concerns, our research undertakes a large-scale quantitative analysis to trace the evolution and distinctiveness of writing styles across disciplines. Specifically, we focus on tracing the evolution of disciplines' writing styles, identifying disciplines' distinct patterns of language usage, and elucidating their differences. Driven by the aims, we attempt to answer the following research questions:

1. Have the writing styles of disciplines changed over an extended period, and how do they evolve?
2. Do the writing styles of different disciplines differ significantly from each other? In other words, are the writing styles of disciplines discriminable?
3. What linguistic features are the most important in differentiating the writing styles of disciplines?

Accordingly, this study makes a three-fold contribution to the literature. A dataset of over 14 million abstracts from the Microsoft Academic Graph (MAG) was used to decode the writing styles of a wide range of disciplines, across eight soft and hard sciences. Additionally, an exhaustive collection of linguistic features was compiled to characterize the writing styles of disciplines. Furthermore, the interpretable machine learning methodology was applied to distinguish the writing styles of the disciplines and unveil the fundamental linguistic features that vary substantially among them. This study reveals the dynamics and variations of scientific writing styles of diverse disciplines in science. The applied computational approach of this study can guide the development of linguistic and text analysis tools for academic publications. Decoding the writing styles of disciplines is crucial for fostering effective interdisciplinary communication and guiding educators in crafting discipline-specific writing curricula.

[1] https://chat.openai.com/.

2. Related work

2.1. Scientific writing for scholarly communication

In comparison to general writing, scientific writing is more specialized, more complex, and briefer (Liu et al., 2022; Tagliacozzo, 1978). It not only conveys the subject matter of research but also reflects disciplinary societal norms, regulations, and common sense (Yore et al., 2004). The fundamental purpose of scientific writing is to communicate inquiries, procedures, and understandings, as well as to facilitate the exchange of ideas among scholars (Hyland, 2004; Ireland & Pennebaker, 2010; Yore et al., 2004).

It has been shown that scholars construct their writings to make their research acceptable to the communities (Hyland, 2004; Sullivan, 1996). Through intentional perspective-taking (Clark & Bernnan, 1991) or automatic interactive alignment (Pickering & Garrod, 2004), readers can acquire knowledge and establish consensus. As well, readers can assess the validity of the asserted claims and make informed decisions on related problems (Norris & Philips, 2003). Scientists who communicate more effectively will receive more recognition and support from the scientific community, leading to better outcomes such as gaining more scientific impact (Ante, 2022; Lu et al., 2019; Gonsalves et al., 2021), more research grants (van den Besselaar & Mom, 2022), better publications (Marino Fages, 2020), and more public attention (Kueffer & Larson, 2014; Jin et al., 2021; Yore et al., 2004).

2.2. Writing style of disciplines

Style can be interpreted as a collection of stylistic markers, including common markers such as diction, syntax, rhetoric, organization of structure, etc. (Lu et al., 2021; Zheng et al., 2006). Linguistic choices shape the thinking of scientists (Langer & Applebee, 1987), improve communication efficiency (Roland, 2009), and contribute to the quality of scientific research (Brown, 2003; Kueffer & Larson, 2014).

The concept of *style* in writing includes *language style* (Ireland & Pennebaker, 2010), *literary style* (Thornborrow & Wareing, 1998), and *linguistic style* (Yang et al., 2021). The term *writing style* emerges as the most commonly employed in scholarly discussions about how content is conveyed in written form (Alluqmani & Shamir, 2018; Ashok et al., 2013; Lu et al., 2019; Song et al., 2023; van den Besselaar & Mom, 2022). *Writing style* specifically refers to the manner in which the content is expressed (Gao et al., 2024; Ireland & Pennebaker, 2010), revealing linguistic characteristics that influence the reader's perception of the text (Liu et al., 2023; Thornborrow & Wareing, 1998).

Writing style can also be regarded as a collective result by scientific communities or journals. Similar language styles can be found in the communities sharing the same education and instruction, in that the scholars may form similar writing habits in word choices and chapter structures (Swales, 1990). Lu et al. (2019) found marginal differences in writing styles between native and non-native English-speaking scholars. As for journals, the articles are subjected to rigorous peer review and proofreading that may strengthen writing styles. As different journals have distinct reviewers and editorial boards with varied requirements and standards, each journal may develop its unique writing style (Nosek et al., 2012; Smaldino et al., 2019). Diachronic changes in scientific writing style have been examined (Sun et al., 2021); Di Feo and Whissell (2020) revealed notable differences in abstract length and inferential statistics between the pre- and post-computer eras according to the abstracts of *the Journal of Consulting and Clinical Psychology*.

Similarly, a discipline is held together by a shared epistemology, an exclusive knowledge structure, and a culture, in which loyalty is generated through unified symbols of the community (Swales, 1990). Its gatekeepers and authorities have developed specific preferences and norms, which might result in a discipline-specific writing style in practice (Choi & Pak, 2008; Frost et al., 1999). A few empirical studies have identified a diversity of linguistic differences among disciplines. Tagliacozzo (1978) examined the differences in writing styles between Genetics and Psychology in terms of the technicality and specialization of language use. Categorical differences in functional word frequencies have been found between historical and experimental sciences (Argamon et al., 2008; Demarest & Sugimoto, 2015). Teich et al. (2016) have confirmed that language use across disciplines has become increasingly different, corroborating the observation that as a discipline develops, it creates its own distinct discourse domain (Halliday & Martin, 2003).

2.3. Quantitative measurements of writing style

Quantitative methods have been utilized to measure linguistic features in many recent studies on authorship identification (Wu et al., 2021; Zheng et al., 2006) and writing style analysis (Ante, 2022; Loughran & Mcdonald, 2014). They cover the five categories: symbolic, lexical, syntactic, structural features, and readability (Song et al., 2023; Wu et al., 2021; Zheng et al., 2006).

Symbolic features are key elements in written language that go beyond alphabetic or word-based content, including punctuation marks, numbers, and special symbols (Zheng et al., 2006). Punctuation marks, in particular, play a vital role in setting the rhythm and tone of sentences (Rangel & Rosso, 2016; Wu et al., 2021). The frequency of digits in a text provides precise measurements and statistical data for substantiating arguments (Zheng et al., 2006).

Lexical features, commonly including total word count, average word length, and type-token ratio, indicate lexical complexity (Eronen et al., 2021; Wu et al., 2021). These features could be categorized into lexical diversity, lexical density, and lexical sophistication (Lu et al., 2019). According to the full-text articles in PLoS, Lu et al. (2019) found a marginal relationship between lexical complexity and scientific impact.

Syntactic features consist of quantitative variables on syntactic forms, reflecting syntactic properties such as complexity and diversity of syntactic forms (Kormos, 2011; Ortega, 2003). Previous studies have shown that syntactic features provide information on sentence organization and help distinguish between authors (Sidorov et al., 2013; Wu et al., 2021). Commonly used features include sentence count, average sentence length, average length of clauses, variation of sentence length, etc. (Kormos, 2011; Lu, 2019; Ojima, 2006). For instance, Song et al. (2023) examined the differences in four syntactic measurements among Information Science and Library Science journals.

Structural features indicate how the author organizes the layout of the article. They are mainly used in short-text authorship tasks, such as programming structure metrics (e.g., placement of comments and use of debugging symbols). De Vel (2000) introduced the number of lines, number of blank lines, number of characters, digit frequency, number of tabs, etc. Wu et al. (2021) parsed part-of-speech, phrase structures, and dependency relationships as features for the task of authorship attribution using NLP tools.

Readability, defined as the ease with which a reader can understand a written text, is a key indicator of text complexity (Klare, 1974; Loughran & McDonald, 2014). A higher score on the readability index indicates a less readable text. This metric has been extensively applied in evaluating the complexity of scientific texts (Ante, 2022; Clatworthy & Jones, 2001; Song et al., 2023). For instance, Lei and Yan (2016) observed an increase in the readability scores of information science abstracts between 2003 and 2012.

From the above studies, a dual research gap can be discerned. Firstly, prior studies have often investigated a couple of disciplines with limited samples. Secondly, studies on disciplinary differences in writing styles lack a complete quantitative linguistic framework, and their adoption of qualitative methods leads to convoluted interpretation of the disparities. In light of these challenges, we suggest an analysis of writing styles across the broader landscape of science by applying a computational approach. We aim to provide a more robust framework for understanding the complexities of writing styles across various fields of science.

3. Methodology

To explore the discernibility of disciplinary writing styles across sciences, this study unfolds in three phases, including tracing the evolution of writing styles across disciplines, identifying distinct linguistic patterns within disciplines, and elucidating pivotal differences. We compiled a large-scale dataset encompassing a range of disciplines from soft to hard sciences (Cole, 1983; Fanelli & Glänzel, 2013) and analyzed writing styles from five aspects: symbolic, lexical, syntactic, structural, and readability. To reveal the evolution trend of scientific writing styles, t tests were conducted to analyze the diachronic changes in writing styles among disciplines. Further, preliminary pairwise *t*-tests were employed to explore the existence of distinct variations in writing styles among various disciplines. This was followed by a comprehensive pairwise writing style classification task, assessing the overall distinguishability of disciplinary writing styles. Finally, we applied a method of interpreting machine learning model to identify crucial linguistic features that differentiate writing styles across disciplinary boundaries.

3.1. Data collection

To investigate the writing styles of the spectrum of science, the hierarchy of science (Cole, 1983) was referenced to determine the

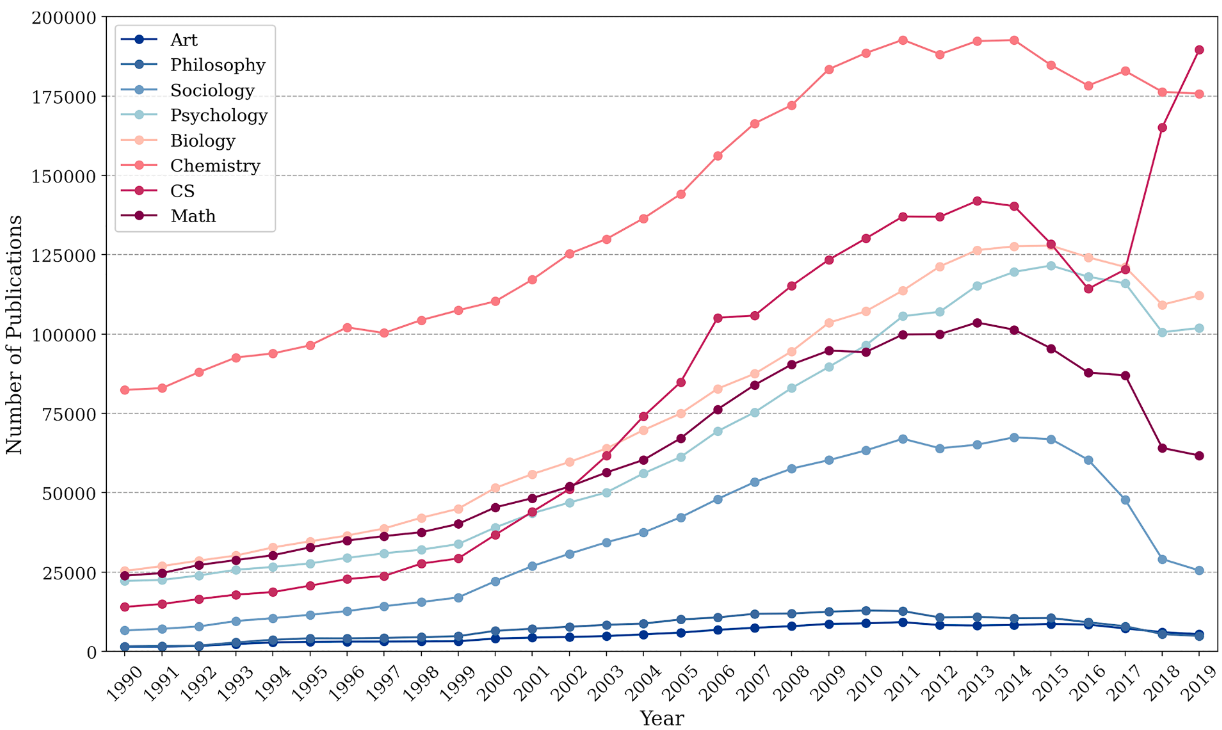

Fig. 1. Number of publications for the eight disciplines from 1990 to 2019.

disciplines to be analyzed. The hierarchy of science presents a holistic view that science is a continuum from soft to hard (Kuhn, 1970). Fanelli and Glänzel (2013) selected six disciplines from soft to hard sciences: Humanities, Social Sciences, Bio-soft Sciences, Bio-hard Sciences, Physical Sciences, and Mathematics. Similarly, we selected eight disciplines with roughly consistent distance intervals from the spectrum of disciplines according to Peng et al. (2021), including Art, Philosophy, Sociology, Psychology, Biology, Chemistry, Computer Science, and Mathematics. Psychology and Biology serve as the demarcation, categorizing the first four as soft sciences and the latter four as hard sciences. It should be noted that this classification primarily indicates tendencies within the disciplines rather than a rigid dichotomy, especially considering that disciplines at the intersection, like Psychology and Biology, frequently display attributes of both categories.

We collected journal abstracts from the Microsoft Academic Graph (MAG, May 2021 version), one of the world's largest academic databases, encompassing publications from a wide array of disciplines. To observe the evolution and divergence of writing styles, we collected journal abstracts from 1990 to 2019 for the eight disciplines. The publications in MAG are assigned with discipline labels (A. K.A. FieldOfStudyIds) and the corresponding confidence scores. A histogram graph of each discipline's confidence scores was formed to determine the peak as the threshold score of the discipline. A publication belongs to a discipline if its confidence score is greater than the threshold score. A total of 14,949,010 publications with at least one discipline assigned were collected for the eight disciplines.

Moreover, publications with missing or excessively short abstracts were excluded. We examined the histogram of abstract lengths across different disciplines, determining the noticeable peak of 40 words as an appropriate minimum length, which precedes the primary distribution peak at around 125 words, especially in Art, Sociology, and Psychology. The choice of this cutoff is significant, as it marks an important juncture in the lower range of abstract lengths across disciplines. Finally, we obtained a dataset with 14,285,952 publications. Fig. 1 shows the annual frequency of the publications for the eight disciplines over the 30 years, illustrating rapid growth in the first two decades, followed by a plateau and a fluctuating decline in most disciplines during the last decade.

3.2. Linguistic features

Quantitative linguistic features are integrated to comprehensively depict the writing styles of disciplines, which are subsequently used as inputs for our machine learning models. The selection of these linguistic features is based on a thorough review of relevant empirical studies in stylometrics (Ante, 2022; Loughran & Mcdonald, 2014) and authorship identification (Wu et al., 2021; Zheng et al., 2006). A comprehensive list of 14 linguistic features were compiled, as detailed in Table 1.

3.2.1. Symbolic features

Digit frequency: the total count of occurrences of the 10 digits (0–9) in the text, providing insights into the utilization of numerical representation (Wu et al., 2021). Digits allow for the precise depiction and comparison of research results, which textual descriptions cannot substitute.

Punctuation frequency: the standard deviation of the frequency of 10 common punctuation symbols (including, *' -; . ! ?% : …*) in the text. Punctuation is auxiliary symbol of written text used to indicate pauses, tone, and function of texts (Rangel & Rosso, 2016; Wu et al., 2021).

Table 1
Definitions of the 14 linguistic features.

Category	Feature	Definition	References
Symbolic	Digit frequency	The count of occurrences of digits (0–9) in the text	Wu et al. (2021); Zheng et al. (2006)
	Punctuation frequency	The standard deviation of the frequency of occurrence of common punctuation symbols in the text	Rangel and Rosso (2016); Wu et al. (2021)
Lexical	Average word length	The average length of words in the text	Stamatatos (2009); van den Besselaar and Mom (2022)
	Number of unique words	The number of words in the text after de-duplication	Crossley (2020); Stamatatos (2009);
	Type-token ratio	The proportion of all different words in the text out of the total number of tokens f	Barrón-Cedeño et al. (2019); Kormos (2011)
	Word count	The total number of words in the text	Di Feo and Whissell (2020); Yang et al. (2021)
Syntactic	Average sentence length	The average length of sentences in the text	Zheng et al. (2006) Jin et al. (2021)
	Sentence count	The total number of sentences in the text	Lu et al. (2021)
	Sentence length dispersion	The variation of sentence length in the text	Xia and Rao (2020)
Structural	Common word ratio	The ratio of the number of common words to the total number of words in the text	Hou et al. (2018)
	POS tag diversity	The diversity of POS tags in the syntactic structure of the text	Ganjigunte Ashok et al. (2013); Fourkioti et al. (2019); Wu et al. (2021)
Readability	Coleman Liau index	Use character, word, and sentence counts to measure readability.	Coleman and Liau (1975)
	Dale-Chall index	Use polysyllables and sentence counts to measure readability.	Chall and Dale (1995); Song et al. (2023)
	SMOG index	Use difficult words and average sentence length to measure readability.	Ante (2022); Jin et al. (2021); McLaughlin (1969)

3.2.2. Lexical features

The features extracted at lexical level include average word length (lexical sophistication), word count (lexical sophistication), number of unique words (lexical diversity), and type-token ratio (lexical density), revealing the lexical writing preferences of disciplines.

Average word length: the average length of words (van den Besselaar & Mom, 2022).

Word count: the total number of words in the text. It reflects the total length of the text in terms of words.

Number of unique words: the number of words after de-duplication, measuring the word types in the text.

Type-token Ratio (TTR): the ratio obtained by dividing the types in the text by its tokens, which is a measure of vocabulary richness (Kormos, 2011; Lu et al., 2019). Types represent all different words, while tokens refer to the total number of words.

3.2.3. Syntactic features

The syntactic features include average sentence length (syntactic length), sentence count (syntactic complexity), and sentence length dispersion (syntactic diversity). These features reflect the length, complexity, and variety of syntactic forms in the text (Kormos, 2011; Ortega, 2003).

Average sentence length: the average length of sentences in the text. Average sentence length is an essential indicator of text's syntactic complexity (Jin et al., 2021; Zheng et al., 2006).

Sentence count: the total number of sentences.

Sentence length dispersion: the variation of sentence length within the text, measured as the standard deviation from the average sentence length (Xia & Rao, 2020). A smaller dispersion indicates a more stable rhythm, while a larger dispersion indicates a more variable rhythm with both fast and slow pacing.

3.2.4. Structural features

Structural features, including Part-of-speech (POS) tag diversity and common word ratio, reflect the structural characteristics of a text (Banerjee et al., 2017; Cao et al., 2011).

POS tag diversity: the diversity of POS tags in the syntactic structure of a given text (Juola, 2008; Wu et al., 2021). This diversity is quantified considering the count of unique POS types, the total count of POS tags, and the entropy of the POS tag frequency distribution. Specifically, the term "(POS tag types) / (POS tag count)" denotes the ratio of unique POS tag types to the total POS tag count, signifying the level of POS type variety. Furthermore, the entropy of the POS tag frequency distribution is used to gauge the uniformity of the POS tag distribution within a text (Ganjigunte Ashok et al., 2013).

$$POS_{Diversity} = \frac{POS\ tag\ types \times entropy\ of\ POS\ tag\ frequency\ distribution}{POS\ tag\ count} \tag{1}$$

Common word ratio: the ratio of the count of common words to the total word count, reflecting the extent to which common words are used in a text (Hou et al., 2018). Common words, composed of systematic functional words and non-semantic stop words, can mirror the structural layout of the paper (Demarest & Sugimoto, 2015; Ireland & Pennebaker, 2010). Our list of common words combines academic keywords (Paquot, 2010) and NLTK stop words (Bird et al., 2009), culminating in a total of 1020 words.

3.2.5. Readability features

Readability is the ease with which a reader can comprehend a written text (DuBay, 2004). Readability indexes provide a comprehensive assessment of a text's readability (Ante, 2022). We selected three indexes, including Coleman Liau, Dale-Chall, and SMOG index.

Coleman-Liau index (Coleman & Liau, 1975) is based on word count, which uses the average sentence length and the average number of characters to measure the readability of a text.

$$Coleman\ Liau = 0.0588 \times \left(\frac{character\ count}{word\ count}\right) - 0.2960 \times \left(\frac{sentence\ count}{word\ count} \times 100\right) - 15.8 \tag{2}$$

Dale-Chall index (Chall & Dale, 1995) uses a list of easy words ($n = 3000$) to measure the vocabulary difficulty of a text. Difficult words were defined as words that were not found in the word list that could be easily understood by 80 of fourth graders. It considers the average sentence length and the percentage of difficult words to measure a text's readability.

$$Dale - Chall = 0.1579 \times \left(\frac{difficult\ words}{total\ words} \times 100\right) + 0.0496 \times \left(\frac{word\ count}{sentence\ count}\right) + 3.6365 \tag{3}$$

SMOG index (McLaughlin, 1969) calculates the readability of a text by counting the number of polysyllables. The higher the value of the index, the more difficult the text is to understand.

$$SMOG = 1.0430 \times \sqrt{\frac{polysyllable\ count \times 30}{sentence\ count}} + 3.1291 \tag{4}$$

3.3. Writing style classification models

To gain a deeper insight into the discernibility of writing styles across disciplines and to explore how disciplinary preferences shape these styles, we designed a task of writing style classification using explainable machine learning approaches. Through classification tasks, we tested whether disciplines exhibit discriminable inherent linguistic characteristics. High classification performance suggests that the writing styles of disciplines in the dataset are distinguishable, while lower performance implies the opposite. Subsequently, feature contributions to the classification were analyzed to decode disciplinary preferences in writing styles. Fig. 2 illustrates the framework of writing style classification.

We formulated the multi-classification of eight disciplines into multiple pairwise tasks. Specifically, 28 pairwise classifiers were created using a subset of abstracts published in 2019. If both disciplines of a classifier belong to soft or hard sciences, it is an in-group comparison; otherwise, an out-group comparison.

Previous studies mainly used SVMs to identify different groups' language use (Argamon et al., 2008; Teich et al., 2016). To leverage the classification results to examine our hypotheses, we compared seven state-of-the-art algorithms, including Decision Tree (DT, Breiman et, al.,1984), Support Vector Machine (SVM, Cortes & Vapnik, 1995), Random Forest (RF, Breiman, 2001), Extreme Gradient Boosting (XGBoost, Chen & Guestrin, 2016), extreme learning machine (ELM, Huang, Zhu, & Siew, 2006), multilayer perceptron (MLP, Rumelhart et al., 1986), and TabNet (Arik & Pfister, 2021).

The first four algorithms, DT, SVM, RF, and XGBoost, are commonly used machine learning algorithms for building classification models with tabular or numerical data. DT is known for its easy understanding (Kotsiantis, 2013), SVM for handling high-dimensional data (Cortes & Vapnik, 1995), RF for its robustness against overfitting (Breiman, 2001), and XGBoost for efficiency and scalability (Chen & Guestrin, 2016). The remaining three, MLP, TabNet, and ELM, are neural networks. MLP excels in generalization (Castillo et al., 2000), TabNet offers interpretability and efficient learning (Shah et al., 2022), and ELM is effective for imbalanced tasks (Huang et al., 2015).

To deal with the imbalanced publications of disciplines (Fig. 1), the Synthetic Minority Over-sampling Technique (SMOTE, Chawla et al., 2002) was applied to oversample the less-publication disciplines. In a trial for Art-Philosophy pair, we found that XGBoost showed higher performance than others (Table 2). Moreover, XGBoost is notably robust in managing feature correlation; for instance, when presented with two perfectly correlated features, it concentrates on one during the boosting process, thereby enhancing the significance of a single feature instead of spreading it across both (Chen & Guestrin, 2016). Therefore, we employed XGBoost to build the final classifiers. The grid search method was employed to determine the optimal hyperparameters.

The correlations between the relative distance between disciplines and the performance of the corresponding classifiers were also investigated. We hypothesize that discipline pairs with greater distances would yield better performance in their pairwise classifications, compared to those with closer distances.

Regarding further feature analysis, we employed the SHapley Additive exPlanations (SHAP, Lundberg & Lee, 2017) to analyze feature contributions. SHAP, based on cooperative game theory (Faigle & Kern, 1992), treats input features as *players* in a game, assigning each feature a Shapley value to quantify its impact on model predictions. This approach calculates contributions based on individual samples, offering a more nuanced understanding of each feature's influence. Mathematically, SHAP values represent the

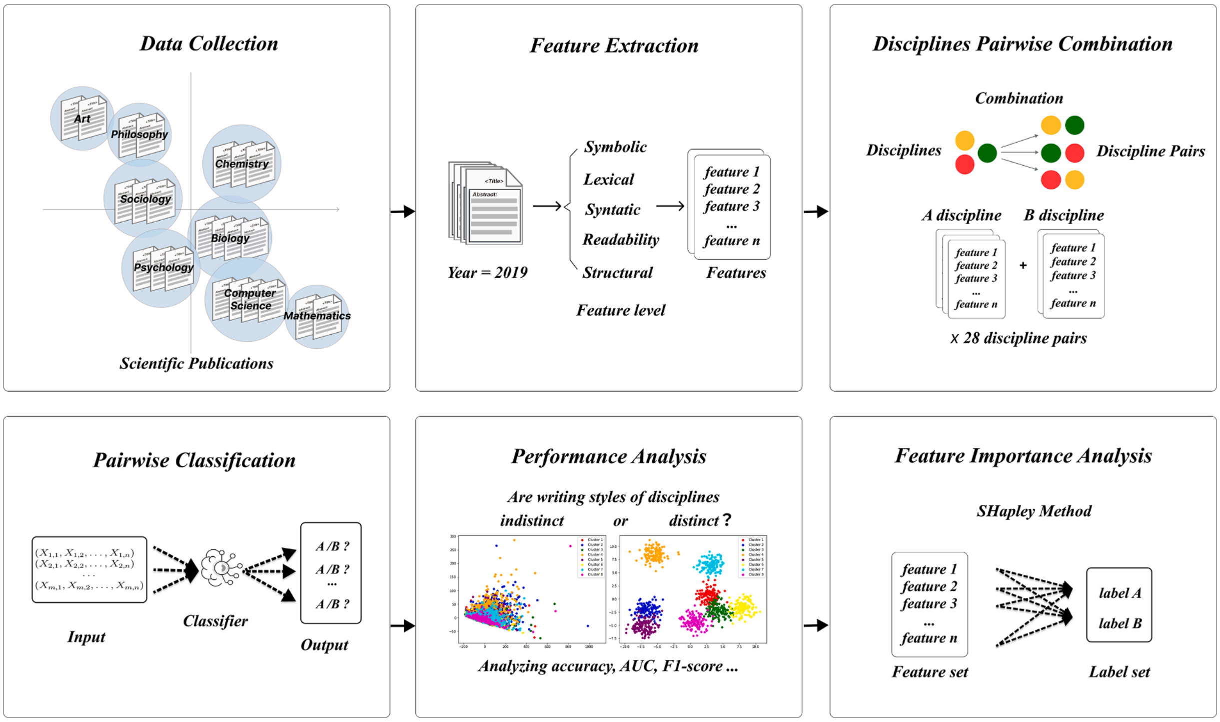

Fig 2. The framework of classifying the writing styles of disciplines and analyzing feature importance.

Table 2
The performance of the four algorithms for Art-Philosophy pair.

Algorithm	Accuracy	Macro-F1	AUC
DT	0.6465	0.6392	0.6386
RF	0.6741	0.6711	0.6689
SVM	0.6707	0.6687	0.6665
XGBoost	**0.6765**	**0.6762**	**0.6747**
ELM	0.5288	0.3505	0.5003
MLP	0.6242	0.6241	0.6722
TabNet	0.6692	0.6631	0.6636

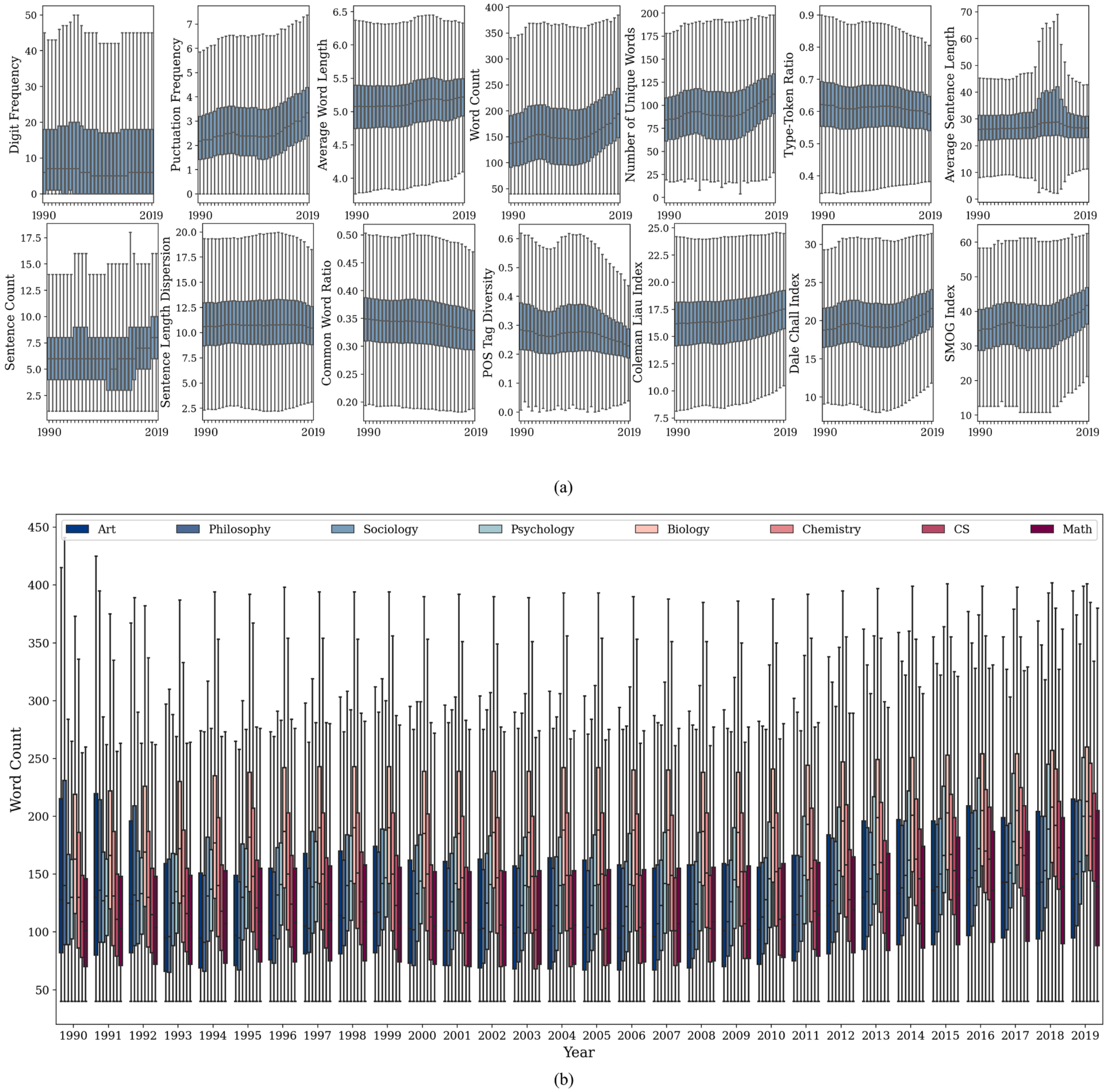

Fig. 3. (a) The evolving trends of 14 linguistic features for all disciplines over 30 years. (b) The evolving trends of the word count for each discipline over 30 years.

allocation of *expenditure* to each feature, considered as contributors in the model. For instance, if x_{ij} is the j^{th} feature of the i^{th} sample, and y_i is the prediction for that sample, SHAP interprets the model as:

$$y_i = y_{base} + f(x_{i1}) + f(x_{i2}) + \ldots + f(x_{ij}) \tag{5}$$

Here, y_{base} is the expected prediction, and $f(x_{ij})$ is the SHAP value, indicating the contribution of each feature. A positive $f(x_{ij})$ suggests a feature enhances the prediction, and vice versa. We measure the global SHAP contribution of each feature by measuring the average absolute value of SHAP across all samples (i.e., paper abstract), shedding light on how differences at the feature level distinguish writing styles across disciplines.

Since we opted for XGBoost, we utilized the TreeExplainer from the SHAP Python package to assess feature contributions. The TreeExplainer is more robust to feature intercorrelations, since it is based on interventional expectations, evaluating each feature's impact independently (Janzing et al., 2020), thereby ensuring more reliable results.

4. Results

4.1. Statistical analysis of diachronic change in writing styles

4.1.1. Diachronic changes in disciplinary writing styles

To explore if scientific discourse has become more complex and specialized, we performed *t*-tests for the first (1990) and last (2019) recorded year in our dataset. Fig. 3a reveals that linguistic features show a relatively consistent trend across all disciplines. Fig. 3b demonstrates the trend of the word count over the 30 years as an example. Trends for all features are depicted in Figs. S1 to S14 of the Supplementary Materials.

In Fig. 3a, lexical complexity exhibits a significant increase. Average word length, word count, and the number of unique words have exhibited fluctuations with an overall upward trend. The first two indicate a growth in lexical sophistication, while the last reflects an increase in lexical diversity. However, there is an exception in the type-token ratio, which shows a decreasing trend, indicating a reduction in lexical density. This decline in the type-token ratio can be attributed to the simultaneous increase in both the type count (number of unique words) and the token count (word count), with the latter experiencing a proportionally greater increase.

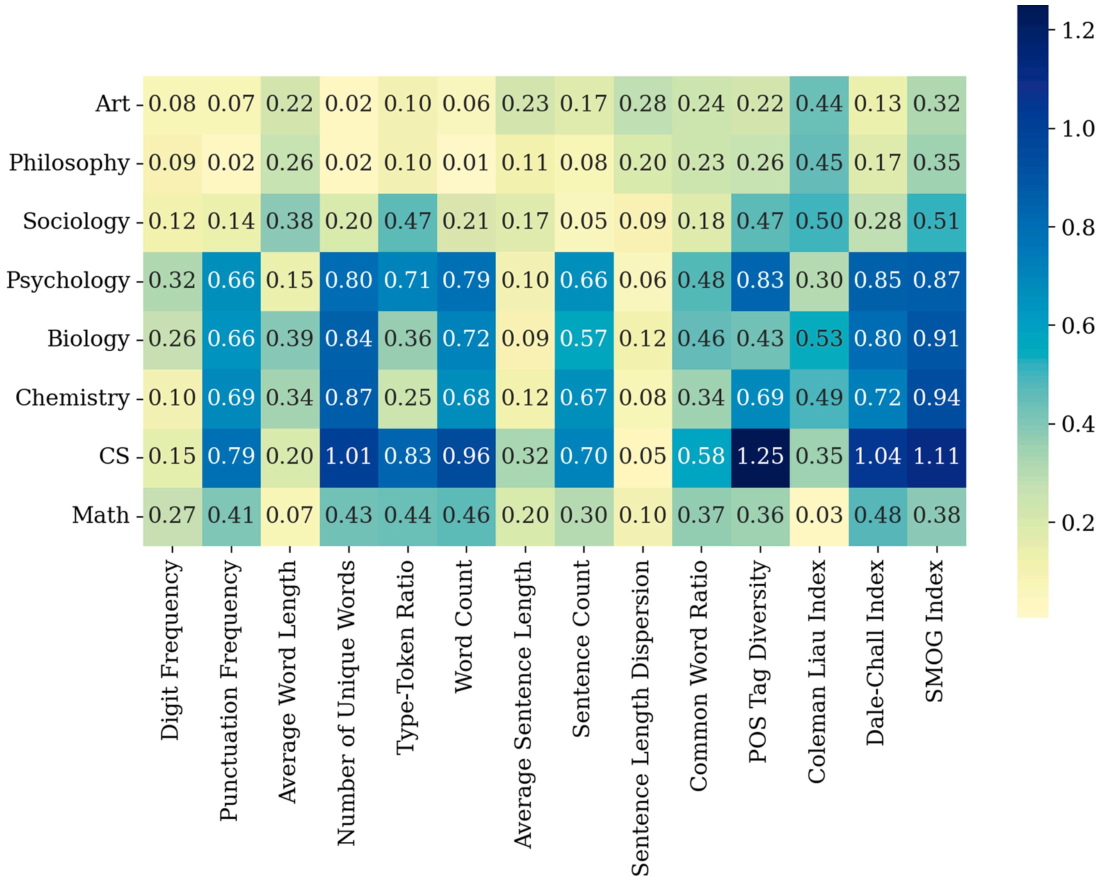

Fig. 4. Cohen's d effect sizes for the 30-year variation of linguist features by disciplines.

Concerning syntactic features, there is a slight increase in syntactic complexity, whereas syntactic diversity and sentence length remain relatively stable. Average sentence length has exhibited fluctuations but has ultimately remained stable, indicating consistent syntactic length. In contrast, sentence count remained relatively stable during the first 20 years but displayed an increasing trend in the last decade, signifying a growing complexity. Sentence length dispersion has remained stable, suggesting a consistent diversity in the narrative rhythm of scientific writing.

Regarding structural features, there is a fluctuating decrease in POS tag diversity, and the common word ratio exhibits a slight declining trend. These observations indicate a reduced use of common words in academic texts and a less diversified structure, suggesting increasing complexity.

Additionally, readability indexes indicate a consistent increase in the Coleman Liau index, with the Dale-Chall index and SMOG index showing fluctuating upward trends. These trends also reflect a growing complexity in written language.

Since the content of abstracts describes research objects and conveys meaningful information (Song et al., 2023), the increase in word count, number of unique words, and sentence count suggests that scientific abstracts have become more informative. Therefore, our results suggest a heightened level of complexity and informativeness in scientific writing. This concurred with the finding in Sun et al. (2021) that scientific writing is evolving towards more specialized terms and harder to read and comprehend.

Further, the Cohen's d effect size was used to measure the magnitude of the difference in the linguistic features of the disciplines between 1990 and 2019. The absolute values of Cohen's d in our comparisons were interpreted by following Cohen's (1992) guidelines, where a value of 0–0.2 indicates a very small effect, 0.2–0.5 a small effect, 0.5–0.8 a medium effect, 0.8–1.2 a large effect, and above 1.2 a very large effect size. Fig. 4 displays the Cohen's d effect sizes of linguistic features' diachronic variations, ranging from 0.01 to 1.25. Notably, three hard sciences (Biology, Chemistry, and Computer Science) and one soft science (Psychology), exhibit high effect sizes, indicating significant changes in their linguistic features. In particular, Psychology exhibits characteristics akin to hard sciences, potentially due to its interdisciplinary nature, encompassing theoretical and applied domains. For these disciplines, a general trend of increasing linguistic complexity is observed with longer texts, a richer vocabulary, and less common words, leading to a tougher challenge of readability (as shown in Tables S4 to S7 of the Supplementary Materials).

Conversely, three soft sciences (Art, Philosophy, and Sociology) and one hard science (Mathematics) display low effect sizes. Mathematics demonstrates more changes compared to the soft sciences, but still considerably fewer changes than other hard sciences. This suggests that the linguistic features of these disciplines have had minor changes in the past 30 years (as shown in Tables S1 to S3 and S8 of the Supplementary Materials). These slightly changed disciplines are generally mature with established scientific practices, methodologies, and ideologies. Overall, hard sciences underwent great changes in writing styles, while soft science remained relatively stable writing styles.

4.1.2. Analysis of disciplinary writing styles

Based on the hierarchy of science theory, our study explores the variations in linguistic features across disciplines. Fanelli and Glänzel (2013) suggests that as one moves from soft to hard sciences, there is a greater scientific consensus, which reduces the effort required for communication (i.e., less need to introduce, justify, and explain), leading to shorter texts, fewer references, etc. Thus, we conjectured that a declining trend in the value of linguistic features from soft to hard sciences would be observed. Contrary to our hypothesis, Fig. 5 reveals an unexpected inverted U-shaped pattern across the disciplines. The values of most linguistic features are minimal at the extremes—Art and Mathematics—and peaks in Biology.

For symbolic features, digit frequency is notably higher in Psychology, Biology, and Chemistry, with Biology and Chemistry showing pronounced usage, indicated by their medians exceeding zero, in contrast to other disciplines. Punctuation frequency exhibits an inverted U-shaped pattern from soft to hard sciences. For lexical features, average word length displays a relatively fluctuated distribution, with shorter lengths in Art and Philosophy. Both unique word count and total word count demonstrate an inverted U-shape, while the Type-Token Ratio (TTR) shows a slight U-shape.

Syntactic features reveal an inverted U-shaped trend in sentence count, with varied sentence length dispersion across disciplines, notably higher in Philosophy and Chemistry. Structural features indicate fluctuating trends in common word ratio and POS tag diversity, with Biology showing lower POS tag diversity. For readability indices, the Coleman-Liau index fluctuates across soft and hard sciences, notably lower in Philosophy and Chemistry, and both the Dale-Chall and SMOG indices exhibit inverted U shapes.

The inverted U-shaped trend may stem from the interdisciplinary nature of central disciplines like Psychology, Biology, and Chemistry. Their integration of methods and terminologies from both soft and hard sciences likely results in more effort to achieve consensus. Furthermore, rapid evolution within their emerging sub-disciplines, synthesizing diverse methodologies, may amplify this need. These heightened features might also reflect the need to communicate with a diverse scholarly audience, as indicated by their extensive publication volume.

4.2. Pairwise comparison between the writing styles of disciplines

T-tests were performed on pairs of disciplines using the data from 2019 to ascertain their discernible differences. Due to the different sample sizes across disciplines (e.g., the largest discipline sample is Chemistry, $n = 240{,}789$, while the smallest discipline sample is Philosophy, $n = 4890$), we performed down sampling for the disciplines except Philosophy ($n = 4890$). Fig. 6 presents the Cohen's d effect sizes for the 28 pairs. Noticeable distinctions are evident in the majority of disciplines, with the Art-related and Philosophy-related pairs displaying particularly pronounced differences and demonstrating similar patterns. Linguistic characteristics exhibited fewer variations among most neighboring disciplines, indicating similar linguistic traits within close disciplines. Additionally, when comparing Mathematics to Biology, Chemistry, CS (hard sciences) and Psychology, consistent patterns of distinction

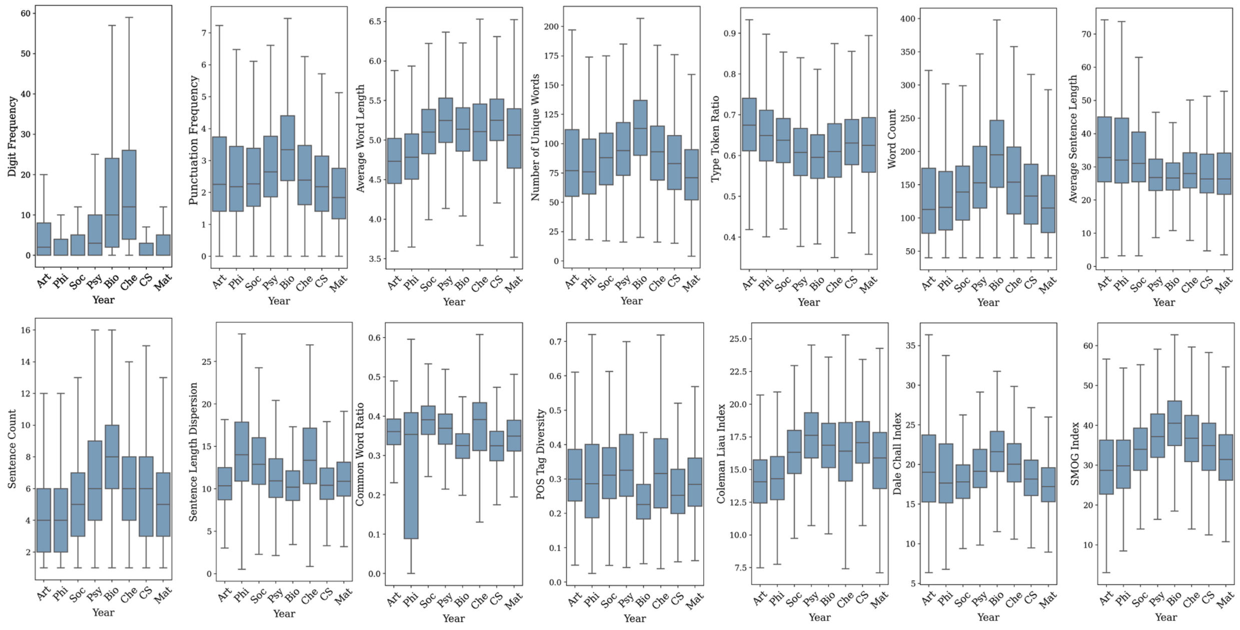

Fig. 5. Boxplots of linguistic features by selected disciplines. * *Each boxplot representing the interquartile range (IQR) of a feature, the median marked by a central line, and 'whiskers' extending to the farthest points within a range of 1.5 times the IQR.*

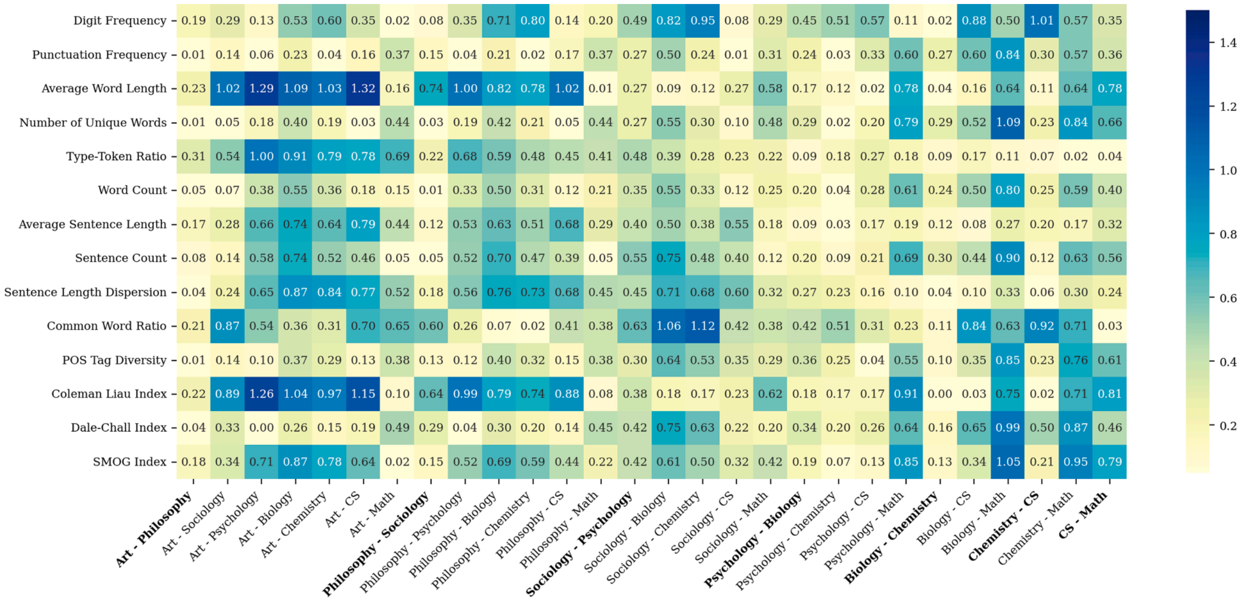

Fig. 6. Cohen's d effect sizes for 28 discipline pairs on 14 linguistic features. Bold x-labels for neighboring discipline pairs.

emerge, manifesting in greater disparities in symbolic features (punctuation frequency), lexical features (average word length, number of unique words, and word count), syntactic feature (sentence count), structure feature (POS tag diversity) and readability (Coleman Liau, Dale-Chall, and SMOG index).

For example, the Psychology-Philosophy comparison (Fig. 7) shows that the two disciplines differ largely in two features (the average word length and Coleman Liau index) and mediumly in five features (type-token ratio, average sentence length, sentence count, sentence length dispersion, and SMOG index). Psychology places higher values on most features than Philosophy. In general, the writing style of Psychology is more complex than that of Philosophy. Further details of these comparisons are available in Tables S9 to S36 of the Supplementary Materials.

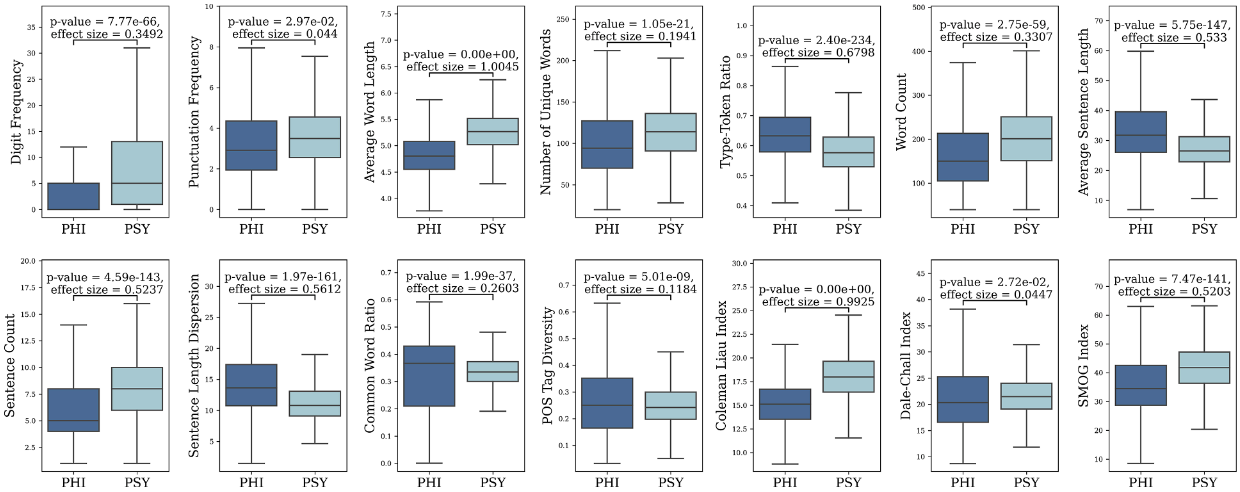

Fig. 7. Boxplots for Philosophy-Psychology Comparison (PHI for Philosophy, PSY for Psychology).

4.3. *Writing styles classification for pairwise disciplines*

Further, machine learning approaches were applied to investigate the discernibility of writing styles between disciplines by constructing classifiers for each pair of disciplines. The classifiers demonstrated good performances: of the twenty-eight classifiers, the accuracy scores of twenty-one exceeded 0.95, four fell between 0.95 and 0.9, two were between 0.8 and 0.9, and one was below 0.8. As depicted in Fig. 8a, the average macro-accuracy per discipline ranged from 0.8515 to 0.9197. A slightly lower average accuracy was observed in disciplines intersecting soft and hard sciences. However, the high accuracies observed in Art- and Philosophy-related classifiers may be attributed to the significantly lower publication quantities in these fields (Fig. 1). This may lead to the model performing well in predicting the majority class but overlooking the minority class.

AUC scores help mitigate the influence of class imbalance and provide a more comprehensive assessment of discriminatory power. The average AUC scores of disciplines vary from 0.7513 to 0.8213, with those related to Chemistry and Art exceeding 0.8, and the remaining six disciplines' scores above 0.75 (Fig 8b). The good performances of the classifiers indicate that the selected linguistic features, coupled with the adopted methods, are effective in differentiating writing styles.

Two interesting observations are drawn from the performances of all the pairwise classifiers as shown in Fig. 9. Firstly, the classifiers along the diagonal line have lower performance, suggesting a trend wherein disciplines in close proximity are more difficult to distinguish. Particularly, Art-Philosophy, Psychology-Sociology, and Biology-Chemistry classifiers yield the lowest AUC scores, indicating challenging discernibility between their writing styles. The difficulties may stem from their close intellectual distance. For instance, both Art and Philosophy rely on discernment, imagination, and perception, have similar research subjects, and focus on

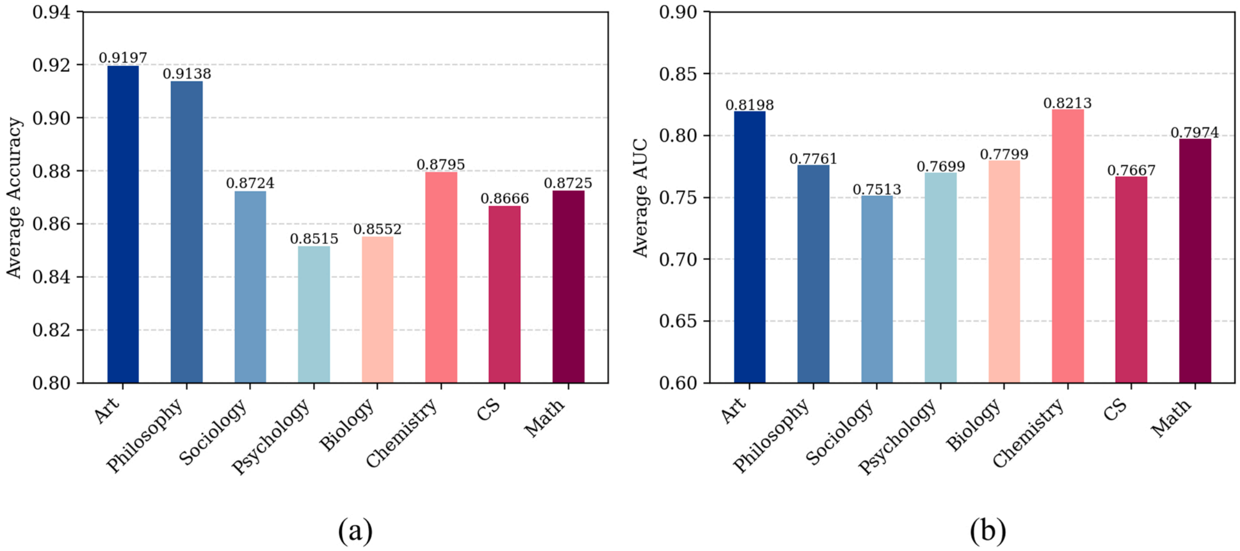

Fig. 8. (a) Average accuracy of classifiers for each discipline. (b) Average AUC of classifiers for each discipline.

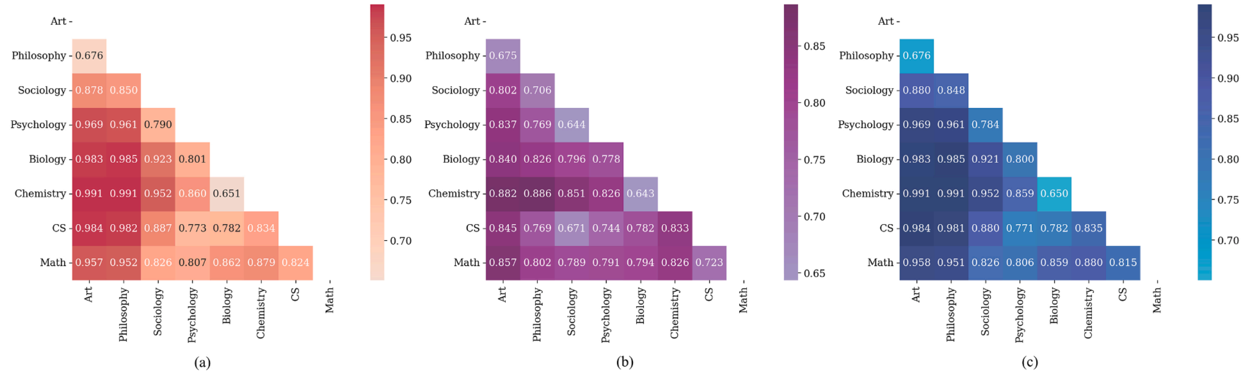

Fig. 9. Performance for each classifier: (a) Accuracy. (b) AUC. (c) Macro-F1 score.

immaterial topics (e.g., sensation, thought, feeling) that lack physical subject constraints. It has been argued that Art and Philosophy are intertwined, just like two sides of the same coin, as “it invites us to intellectual consideration, and that not for the purpose of creating art again, but for knowing philosophically what art is.” (Hegel, 1975, p.11). Since disciplinary practices are constrained by their research subjects (Fanelli & Glänzel, 2013), the commonality in research subjects and methodology of Art and Philosophy may lead to their writing styles being challenging to distinguish. On contrary, the longest distance between disciplines would yield the best classification performance, Art and Mathematics, two poles of the soft and hard spectrum, have a high AUC of 0.857 (ranks 3/28).

To gain further insight into the relationship between the relative distance of disciplines and their pairwise classification performance, we conducted a correlation analysis. As illustrated in Fig. 10a, the correlation coefficient between relative distance and AUC for all comparisons is 0.505, indicating a significant positive correlation (p-value $= 0.0062 < 0.01$). Thus, the greater the distance between the two disciplines, the more their writing styles tend to differ. This positive correlation is particularly pronounced in-group comparisons, showing a coefficient of 0.679 (p-value $= 0.0153$), while less prominent in out-group comparisons (a coefficient of 0.227, p-value $= 0.3971$).

Moreover, a correlation analysis was conducted between the relative discipline distance and the average effect size of disciplines (Fig. 10b). A higher average effect size signifies superior comprehensive discernability. The correlation between the average effect sizes and the relative disciplinary distances for all comparison pairs is not readily discernible, as demonstrated by a weak correlation coefficient of 0.183 (p-value $= 0.3525$). However, a notable positive correlation between average effect size and distance appears in-group pairs (coefficient $r = 0.708$, p-value $= 0.0099$), but only a weak correlation appears in out-group comparisons (coefficient $r = 0.378$, p-value $= 0.1483$). These observations align with the AUC correlation results, demonstrating that across various scientific fields, the discernibility in writing styles increases with the disciplinary distance, regardless of the measurement method used.

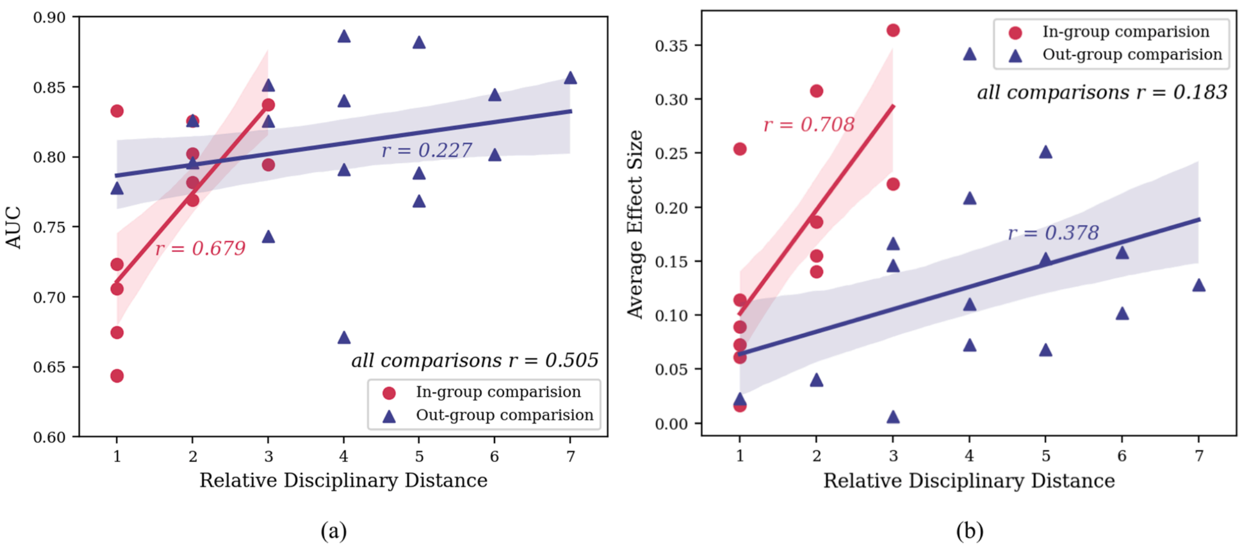

Fig. 10. (a) The distribution of classification accuracy over relative disciplinary distance, with a Pearson correlation coefficient of 0.505. (b) The distribution of the average effect size over the relative disciplinary distance between disciplines, with a Pearson correlation coefficient of 0.183. * *Shaded regions represent confidence intervals with CI = 0.8.*

4.4. Contribution analysis of linguistic features

To further decode writing style differences, we analyzed the differences in linguistic features between disciplines by interpreting the contributions of linguistic features to the pairwise classifications.

4.4.1. Feature ranking

Ordered by the SHAP values of features, the average ranks of features were calculated in the four sets of classifiers, namely, (1) all 28 classifiers, (2) in-group classifiers of soft sciences, (3) in-group classifiers of hard sciences, and (4) out-group classifiers between soft sciences and hard sciences. As shown in Table 3, the importance of feature contribution remains generally consistent across classifiers, and all features maintain relatively stable ranking positions, with fluctuations of no more than 2.19 positions compared to the average rank. This consistency holds true for in-group comparisons of both soft and hard sciences, as well as out-group comparisons.

Using the average ranks to assess feature importance, it is evident that certain features consistently play a more significant role in classifying different disciplines. Symbolic, lexical, and structural features have a feature ranked among the top seven (half of the 14 numeric features): digit frequency, average word length, and common word ratio. Each readability index is ranked among the top five, serving as an essential feature for the writing style classification task. While, syntactic features and other features rank at the bottom, indicating a minimal contribution to the classification.

The variations in the ranks of critical features across the four sets of classifiers highlight differences between soft and hard sciences. For in-group classifiers of hard sciences, the role of digit frequency in distinguishing writing styles is substantial, as evidenced by a 0.29 position rise from an average rank of 6.79 to 6.5. This implies that digit utilization serves as a key differentiator within hard sciences. While, this factor is not as impactful in soft sciences, where it maintains an average rank of 8.5, so it also serves as a strong factor in discerning between soft and hard sciences (ranks up to 6.25). Moreover, the average word length ranks 3.83 in-group comparisons of hard sciences, notably higher than the overall average rank of 5.61.

For in-group classifiers of soft sciences, the average sentence length and the Coleman-Liau index significantly contribute to the differentiation. Among in-group classifiers of soft sciences, average sentence length ranks 6.83, while it does not appear in the top 7 in other classifiers. This suggests that average sentence length is not a vital feature in other classifiers, but it forms meaningful classification rules when differentiating between soft sciences. In other words, the average sentence length among soft sciences is significantly different. In in-group classifiers for soft sciences, the Coleman-Liau index increased its rank position by 2.18, emphasizing the importance of the readability based on word and sentence count for classification.

4.4.2. Linguistic feature contribution patterns

To further investigate the intricate feature contribution patterns, we scrutinized the feature contribution across 28 discipline-specific classifiers (as shown in Figs. S15 to S22 of the Supplementary Materials). We utilized Kendall's Tau rank correlation coefficient to measure the correlation between every two lists of feature rankings (arranged according to the features in Table 1) from different classifiers. For every two disciplines, there are two sets of rankings, and every set consists of six rankings from the six classifiers. We computed the Kendall's Tau rank correlation coefficients between the feature contribution rankings of the classifiers, then calculated the average of the six coefficients to represent the resemblance of feature contribution pattern between disciplines. And we calculated the average rank coefficient for each discipline against all soft sciences and all hard sciences separately.

Table 4 shows that most pairs of disciplines show a low average rank correlation around 0.0623 to 0.3004, except for a much higher

Table 3
The average ranking of numeric features for the four sets of classifiers.

Category	Features	Avg. rank of all classifiers	Avg. rank of in-group classifiers for soft sciences	Avg. rank of in-group classifiers for hard sciences	Avg. rank of out-group classifiers
Symbolic	Digit Frequency	**6.79**	6.79+1.71	**6.79–0.29**	**6.79–0.54**
	Punctuation Frequency	9.04	9.04+1.13	9.04–1.37	9.04+0.09
Lexical	Average Word Length	**5.61**	**5.61–0.44**	**5.61–1.78**	**5.61+0.83**
	Number of Unique Words	7.64	7.64+2.19	7.64–0.14	**7.64–0.76**
	Type-Token Ratio	9.86	9.86–0.03	9.86–1.19	9.86+0.45
	Word Count	9.32	9.32+2.01	9.32+1.85	9.32–1.45
Syntactic	Average Sentence Length	8.57	**8.57–1.74**	8.57+0.26	8.57+0.55
	Sentence Count	9.75	9.75–1.42	9.75–0.42	9.75+0.69
	Sentence Length Dispersion	9.04	9.04–1.54	9.04–1.21	9.04+1.02
Structural	POS Tag Diversity	9.54	9.54–1.71	9.54+0.13	9.54+0.59
	Common Word Ratio	**6.89**	**6.89–0.39**	**6.89–0.39**	6.89+0.3
Readability	Coleman Liau Index	**4.18**	**4.18–2.18**	**4.18+1.65**	**4.18+0.2**
	Dale-Chall Index	**4.96**	**4.96+1.21**	**4.96+0.87**	**4.96–0.77**
	SMOG Index	**3.82**	**3.82+1.18**	**3.82+2.01**	**3.82–1.19**

*Bolded figures indicate that the average rank of the feature is within top seven.

correlation found between Biology and Chemistry with an average coefficient of 0.5201. The generally low rank correlation among most discipline pairs highlights the distinct feature contribution patterns unique to each discipline, underscoring the complexity and diversity of their methodology and norms. However, within the same category (either hard or soft sciences), we observed higher coefficients, indicating the complicated interrelationships and tacit norms behind same category disciplines. This result may help to explain the relatively lower performance in-group classifications. Moreover, it is found that Biology-related and Chemistry-related classifiers show very similar patterns of feature contribution, which echoes with our previous observation that Biology-Chemistry classifier gets the lowest performance (Fig. 9), providing a reason for this worst performance.

5. Discussions

Based on the aforementioned findings, we discuss the theoretical contributions of this study, including evolution patterns of writing styles in science, disparities among disciplines, and crucial linguistical features that dominates the discrimination of disciplines' writing styles. In addition, we deliberate potential practical significances and the limitations of this study.

5.1. Theoretical implications

5.1.1. Evolution patterns of writing styles in science

We investigated eight disciplines, covering a broad scope of soft and hard sciences. Our quantitative analysis on the large-scale publication abstracts discloses that the linguistic features of soft disciplines (Art, Philosophy, and Sociology) and Mathematics generally remained stabilized, experiencing modest change over the 30 years. In contrast, those of the hard sciences (Biology, Chemistry, and Computer Science) are still developing. The difference may be due to the relatively stable research paradigms in the soft sciences, while research methods and objects in the hard sciences are constantly changing, leading to the change in writing styles.

As for those linguistic-dynamic disciplines, including Biology, Chemistry, Computer Science, and Psychology, their writing styles trend towards greater complexity and information richness. There has been an increase in the amount and complexity of information in abstracts due to the growth in vocabulary, use of professional terminology, and abstract length. We speculate that this inclination may be associated with the intensive academic competition, which drives an increase in research workload and clarity of scientific writing, particularly in top-tier journals such as Nature, Science, and Cell (Shi et al., 2023). In addition, some writing traits have experienced minimal shifts, including the use of digits and the sentence length. It reflects the idea of standardization in scientific writing, which entails stringent standards and formatting requirements to maintain clarity, precision, and consistency. The stable use of digits suggests that the academy has obeyed a consistent norm of using data to report their results in abstracts. The balanced use of long and short sentences in scientific writing can better convey complex ideas and concepts. Overly long sentences may lead to difficulties in comprehension for the reader, while excessively short sentences may not fully express intricate viewpoints or results.

From the perspective of the hierarchy of science, a U-shape trend of initially increasing and then decreasing linguistic complexity from soft sciences to hard sciences has been observed (Fig. 5). The initial upward trend can be ascribed to the increasing experimental and professionalized research methodologies and use of discipline-specific terminologies in soft sciences, resulting in a notable increase in complexity and workload. Then, disciplines lying at the intersection of soft and hard sciences, like Psychology, Biology, and Chemistry, exhibit the highest complexity, this may due to their synthesis of diverse methodologies and terminologies and the need to communicate with a broad scholarly audience. Harder sciences have already widely adopted terminologies and methodologies, benefitting from a more well-defined common ground. Consequently, the observed decrease in complexity can be attributed to the reduced need for scholars in harder sciences to explain their research, resulting in more succinct and straightforward writing styles.

5.1.2. The discernable writing styles between disciplines

In addition to basic statistic method (i.e., the pairwise *t*-test), we adopted the machine learning classification to investigate the discriminability of writing styles between disciplines. The good performance of classifiers shows that even without semantics, linguistic features exhibit excellent performance in classifying the publications from a broad scope of disciplines. It can be inferred that all the disciplines may demonstrate their own writing styles. The cultures and conventions manifested by the writing styles are significantly different between disciplines. According to Crossley (2020) and Song et al. (2023), disparities in language usage exist between "micro-disciplines". Our research extends this finding to "macro-disciplines" with strong evidence of disciplinary linguistic differences. Consequently, research that relies on text analysis for deriving insights may need to critically reassess the substantial contributions of textual content, acknowledging that latent stylistic elements may significantly influence the outcomes. This consideration is crucial across a broad spectrum of text-analytic studies, beyond the specific scient metrics contexts.

Our results also indicate that scientific writing is disciplinary-oriented and subservient that scientific writing must incorporate scientific research with the traditions and conventions of the respective discourse community. This confirms the idea of "discourse community" proposed by Swales (1990) that disciplines established their own writing regulation and cultures to facilitate communication within academic communities, and provides convincing findings against "universalism" that scientific papers are decontextualized and autonomous without a shared situational context (DeVito, 1966).

Additionally, the different performances of the classifiers indicate that the greater the distance between disciplines, the more pronounced the differences in language style. This finding highlights the lesser discernibility of writing styles in neighboring disciplines compared to others. The observed high correlation between the discernibility of writing styles and disciplinary distance, further illustrates that the discernibility of writing styles reflects the cultural and epistemic dissimilarity between disciplines.

Table 4
Average Kendall's Tau rank correlation coefficient between disciplines.

Disciplines	Art	Philosophy	Sociology	Psychology	Biology	Chemistry	CS	Math	Avg. coef with soft sciences	Avg. coef with hard sciences
Art		0.1795	0.2418	0.3004	0.1319	0.0842	0.2601	0.1026	0.2406	0.1447
Philosophy			0.2125	0.1758	0.1612	0.2088	0.2088	0.1392	0.1893	0.1795
Sociology				0.2894	0.2088	0.0952	0.2564	0.0623	0.2479	0.1557
Psychology					0.1868	0.0952	0.2711	0.0842	0.2552	0.1593
Biology						**0.5201**	0.2198	0.2308	0.1722	0.3236
Chemistry							0.2234	0.1685	0.1209	0.3040
CS								0.1392	0.2491	0.1941
Math									0.0971	0.1795

*Bolded figure indicates the largest average Kendall's Tau rank correlation coefficient between disciplines.

5.1.3. Important linguistic features for the discrimination of writing styles between disciplines

The interpretable machine learning approach adopted in this study facilitates the analysis of important linguistic features that can discriminate the writing styles between disciplines. All features maintain consistent ranking positions across all classifiers, with readability indexes, digit frequency, average word length, and common word ratio consistently making significant contributions to writing style classifications, while syntactic features and rest features have less contribution to classifications. These results suggest that quantitative features pertaining to the richness and complexity of writing may serve as a strong indicator for classification, that features reflecting the overall readability of discourse (readability indexes) and word complexity (average word length), information richness (common word ratio), and the use of digits (digit frequency) contribute significantly to the classification of distinct writing styles.

Moreover, we found that the feature contribution patterns differ between hard and soft disciplines, as hard sciences show a more distinct use of digits. In contrast, for soft sciences, the average sentence length contributes more significantly, underscoring the value of syntactic complexity in distinguishing soft sciences. Furthermore, the feature contribution analysis enhances the explainability of the classifications. The revealed higher feature rank correlations within the same category help to elucidate the lower performance of in-group classifiers, exemplified by the Biology-Chemistry classifier.

5.2. Practical implications

Technically, the good performance of the machine learning models that incorporates the well-designed framework of linguistic features, indicates that the methodology could guide the development of NLP tools with various applications. A tool based on the implementation of disciplinary writing style classifiers can help researchers check the consistency of their writing with the common style of the target discipline. Similarly, it is also potential to develop methods of detecting AI generated academic content, e.g., by ChatGPT, using writing styles. Publishers could also develop journal-specific writing style classifiers by applying our methodology.

Our findings offer a quantified analysis of writing patterns within specific disciplines, thereby serving as a valuable resource for novices and researchers venturing into unfamiliar fields or collaborating with experts from different domains. Educationally, our work can assist librarians, educators, and publishers in creating targeted writing manuals for novices, compared with typical ones that provide only general recommendations for abstract length (e.g., UW-Madison's guideline suggests 6–7 sentences or 150–250 words[2]). Moreover, taking a diachronic perspective, our findings show that scientific writing styles evolve alongside disciplinary and societal shifts, highlighting the necessity of ongoing research into their evolution. This knowledge is vital for developing writing programs that adapt to the altering language needs of scientific fields.

5.3. Limitations

While our study has provided valuable insights, we acknowledge certain limitations. First, pairwise comparisons do not provide a comprehensive overview of differences across all disciplines, leaving room for further improvements in our methods. Second, we only focused on selected disciplines; further research across a broader range of disciplines could generalize our results. Third, it should be noted that our results are only drawn from publication abstracts, which reflects one significant aspect of scientific discourses. However, the writing styles embedded in the full texts of publications are much more complex with more structural and linguistic characteristics at the article level. Applying machine learning methods to detect the differences in writing styles between disciplines, and develop useful NLP tools for scientific community, will be a challenging study for future. Lastly, many of the quantitative linguistic features used in this paper are superficial. Thus, more advanced indicators that can capture in-depth nuances of writing styles could be developed in further studies.

[2] https://writing.wisc.edu/handbook/assignments/writing-an-abstract-for-your-research-paper/.

6. Conclusion

This study utilized computational linguistics and interpretable machine learning to classify a large-scale bibliographic dataset to investigate the writing styles of disciplines across a broad landscape of science. Our findings reveal an evolving trend towards more complexity and information richness in progressing disciplines' writing language, with mature disciplines' writing styles having stabilized. The high AUC scores of the pairwise writing style classifiers indicate a well discriminability of the writing styles between disciplines. A correlation between the classification AUC and the distance between disciplines was identified. Additionally, discipline-specific patterns of linguistic feature differentiation were discovered to illustrate the characteristics of each discipline, with commonalities found in soft sciences and dissimilarities found in hard sciences.

Overall, this study serves as a starting point for exploring writing styles in scientific publications, potentially offering a novel framework for evaluating significant milestones in the evolution of scientific disciplines, especially in the context of scientific writing styles. We plan to delve deeper into the linguistic interpretations of interdisciplinary domains, shedding light on the evolution of writing styles in emerging interdisciplinary fields in relation to established disciplines.

CRediT authorship contribution statement

Shuyi Dong: Writing – original draft, Visualization, Validation, Methodology, Data curation. **Jin Mao:** Writing – original draft, Methodology, Investigation, Conceptualization. **Qing Ke:** Writing – review & editing, Investigation. **Lei Pei:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors have no competing interests to declare.

Data availability

The authors do not have permission to share data.

Acknowledgement

We thank the anonymous reviewers for their constructive comments that improve the article significantly. We greatly appreciate the help from Zhentao Liang on the data collection and preprocessing. This study was funded by the National Natural Science Foundation of China (NSFC) Grant Nos. 71921002 and 72174154.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ipm.2024.103718.

References

Abelson, P. H. (1980). Scientific Communication. *Science (New York, N.Y.), 209*(4452), 60–62. https://doi.org/10.1126/science.7280664

Alluqmani, A., & Shamir, L. (2018). Writing styles in different scientific disciplines: A data science approach. *Scientometrics, 115*(2), 1071–1085.

Ante, L. (2022). The relationship between readability and scientific impact: Evidence from emerging technology discourses. *Journal of Informetrics, 16*(1), Article 101252. https://doi.org/10.1016/j.joi.2022.101252

Argamon, S., Dodick, J., & Chase, P. (2008). Language use reflects scientific methodology: A corpus-based study of peer-reviewed journal articles. *Scientometrics, 75*(2), 203–238. https://doi.org/10.1007/s11192-007-1768-y

Arik, S.Ö., & Pfister, T. (2021). TabNet: Attentive Interpretable Tabular Learning. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(8), 6679–6687. https://doi.org/10.1609/aaai.v35i8.16826

Ashok, V. G., Feng, S., & Choi, Y. (2013). Success with style: Using writing style to predict the success of novels. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 1753–1764).

Banerjee, S., Chua, A. Y. K., & Kim, J. J. (2017). Don't be deceived: Using linguistic analysis to learn how to discern online review authenticity. *Journal of the Association for Information Science and Technology, 68*(6), 1525–1538. https://doi.org/10.1002/asi.23784

Barrón-Cedeño, A., Jaradat, I., Da San Martino, G., & Nakov, P. (2019). Proppy: Organizing the news based on their propagandistic content. *Information Processing & Management, 56*(5), 1849–1864. https://doi.org/10.1016/j.ipm.2019.03.005

Berkenkotter, C., Huckin, T., & Ackerman, J. (1991). Social context and socially constructed texts: The initiation of a graduate student into a writing research community. In C. Bazerman, & J. Paradis (Eds.), *Textual dynamics of the professions* (pp. 191–215). Madison, WI: University of Wisconsin Press.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python: Analyzing text with the natural language toolkit.* O'Reilly Media, Inc.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and regression trees, 37* pp. 237–251). Wadsworth Int. Group.

Breiman, L. (2001). Random forests. *Machine learning, 45*, 5–32.

Brown, T. L. (2003). *Making truth: Metaphor in science*. University of Illinois Press.

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems, 50*(2), 511–521. https://doi.org/10.1016/j.dss.2010.11.009

Castillo, P. A., Carpio, J., Merelo, J. J., Prieto, A., Rivas, V., & Romero, G. (2000). Evolving multilayer perceptrons. *Neural Processing Letters, 12*(2), 115–128.

Chall, J. S., & Dale, E. (1995). *Readability revisited, the new dale-chall readability formula*. Cambridge, MA: Brookline Books.
Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357.
Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794). pp.
Choi, B. C. K., & Pak, A. W (2008). *Multidisciplinarity, interdisciplinarity, and transdisciplinarity in health research, services, education and policy: 3. discipline, inter-discipline distance, and selection of discipline* (pp. E41–E48). Clinical and Investigative Medicine. https://doi.org/10.25011/cim.v31i1.3140
Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *In Perspectives on socially shared cognition* (pp. 127–149). American Psychological Association. https://doi.org/10.1037/10096-006. pp.
Clatworthy, M., & Jones, M. J. (2001). The effect of thematic structure on the variability of annual report readability. *Accounting, Auditing & Accountability Journal, 14* (3), 311–326. https://doi.org/10.1108/09513570110399890
Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science, 1*(3), 98–101.
Cole, S. (1983). The Hierarchy of the Sciences? *American Journal of Sociology, 89*(1), 111–139. https://doi.org/10.1086/227835
Coleman, M., & Liau, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology, 60*(2), 283.
Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1007/BF00994018
Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research, 11*(3), 3. https://doi.org/10.17239/jowr-2020.11.03.01. Article.
de Vel, O. (2000). Mining e-mail authorship. In Proc. Workshop on Text Mining, ACM International Conference on Knowledge Discovery and Data Mining (KDD'2000).
Demarest, B., & Sugimoto, C. R. (2015). Argue, observe, assess: Measuring disciplinary identities and differences through socio-epistemic discourse. *Journal of the Association for Information Science and Technology, 66*(7), 1374–1387. https://doi.org/10.1002/asi.23271
DeVito, J. A. (1966). Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech Monographs, 33*(1), 73–76.
Di Feo, G., & Whissell, C. (2020). Changes in the Style and Contents of Abstracts from The Journal of Consulting and Clinical Psychology between the 1960s and the 2010s. *Advances in Social Sciences Research Journal, 7*(10), 546–551. https://doi.org/10.14738/assrj.710.9215
DuBay, W.H. (2004). The Principles of Readability. *Online Submission*. https://eric.ed.gov/?id=ed490073.
Eronen, J., Ptaszynski, M., Masui, F., Smywiński-Pohl, A., Leliwa, G., & Wroczynski, M. (2021). Improving classifier training efficiency for automatic cyberbullying detection with Feature Density. *Information Processing & Management, 58*(5), Article 102616. https://doi.org/10.1016/j.ipm.2021.102616
Faigle, U., & Kern, W. (1992). The Shapley value for cooperative games under precedence constraints. *International Journal of Game Theory, 21*, 249–266.
Fanelli, D., & Glänzel, W. (2013). Bibliometric evidence for a hierarchy of the sciences. *PloS one, 8*(6), e66938. https://doi.org/10.1371/journal.pone.0066938
Fourkioti, O., Symeonidis, S., & Arampatzis, A. (2019). Language models and fusion for authorship attribution. *Information Processing & Management, 56*(6), Article 102061. https://doi.org/10.1016/j.ipm.2019.102061
Frost, S.H., Jean, P.M., & Vaughn, V. (1999). *Distances between disciplines: Influences of interdisciplinary discourse on faculty scholarship and interaction at one university*. 36.
Ganjigunte Ashok, V., Feng, S., & Choi, Y. (2013). Success with style: Using writing style to predict the success of novels. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1753–1764). https://aclanthology.org/D13-1181.
Gao, Y., Liu, Q., Yang, Y., & Wang, K. (2024). Latent representation discretization for unsupervised text style generation. *Information Processing & Management, 61*(3), Article 103643. https://doi.org/10.1016/j.ipm.2024.103643
Gonsalves, C., Ludwig, S., de Ruyter, K., & Humphreys, A. (2021). Writing for impact in service research. *Journal of Service Research, 24*(4), 480–499. https://doi.org/10.1177/10946705211024732
Halliday, M.A.K. (1978). *Language as social semiotic*.
Halliday, M. A. K., & Martin, J. R. (2003). *Writing science: Literacy and discursive power*. Taylor & Francis.
Harrison, T., & Stephen, T. (1995). The electronic journal as the heart of an online scholarly community. *Library Trends, 43*(4), 592–608.
Hegel, G. W. F. (1975). *Aesthetics: Lectures on fine art*. Clarendon Press.
Hou, L., Li, J., Li, X. L., Tang, J., & Guo, X. (2018). Learning to align comments to news topics. *ACM Transactions on Information Systems, 36*(1), 1–31. https://doi.org/10.1145/3072591
Huang, G., Huang, G. B., Song, S., & You, K. (2015). Trends in extreme learning machines: A review. *Neural Networks, 61*, 32–48.
Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: theory and applications. *Neurocomputing, 70*(1–3), 489–501.
Hyland, K. (2004). *Disciplinary discourses, michigan classics ed.: Social interactions in academic writing*. University of Michigan Press. https://doi.org/10.3998/mpub.6719
Hyland, K., & Bondi, M. (2006). Academic discourse across disciplines. *Peter Lang CH*. https://doi.org/10.3726/978-3-0351-0446-2
Iaccarino, M. (2003). Science and culture. *EMBO Reports, 4*(3), 220–223. https://doi.org/10.1038/sj.embor.embor781
Ireland, M. E., & Pennebaker, J. W. (2010). Language style matching in writing: Synchrony in essays, correspondence, and poetry. *Journal of Personality and Social Psychology, 99*(3), 549–571. https://doi.org/10.1037/a0020386
Janzing, D., Minorics, L., & Blöbaum, P. (2020). Feature relevance quantification in explainable AI: A causal problem. In *In International Conference on artificial intelligence and statistics* (pp. 2907–2916). PMLR. pp.
Jin, T., Duan, H., Lu, X., Ni, J., & Guo, K. (2021). Do research articles with more readable abstracts receive higher online attention? Evidence from Science. *Scientometrics, 126*(10), 8471–8490. https://doi.org/10.1007/s11192-021-04112-9
Juola, P. (2008). Authorship attribution. *Foundations and Trends® in Information Retrieval, 1*(3), 233–334.
Kaslow, N. J. (2015). Translating psychological science to the public. *American Psychologist, 70*, 361–371. https://doi.org/10.1037/a0039448
Klare, G. R. (1974). Assessing readability. *Reading Research Quarterly, 10*(1), 62–102. https://doi.org/10.2307/747086
Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20*(2), 148–161. https://doi.org/10.1016/j.jslw.2011.02.001
Kotsiantis, S. B. (2013). Decision trees: A recent overview. *Artificial Intelligence Review, 39*, 261–283.
Kueffer, C., & Larson, B. M. H. (2014). Responsible use of language in scientific writing and science communication. *Bioscience, 64*(8), 719–724. https://doi.org/10.1093/biosci/biu084
Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd edition). The University of Chicago Press.
Langer, J. A., & Applebee, A. N. (1987). *How writing shapes thinking: A study of teaching and learning*. National Council of Teachers of English.
Larivière, V., Haustein, S., & Börner, K. (2015). Long-distance interdisciplinarity leads to higher scientific impact. *PloS one, 10*(3), Article e0122565. https://doi.org/10.1371/journal.pone.0122565
Lei, L., & Yan, S. (2016). Readability and citations in information science: Evidence from abstracts and articles of four journals (2003–2012). *Scientometrics, 108*(3), 1155–1169. https://doi.org/10.1007/s11192-016-2036-9
Li, K., & Jiao, C. (2022). The data paper as a sociolinguistic epistemic object: A content analysis on the rhetorical moves used in data paper abstracts. *Journal of the Association for Information Science and Technology, 73*(6), 834–846. https://doi.org/10.1002/asi.24585
Liu, Y., Medlar, A., & Głowacka, D. (2022). Lexical ambiguity detection in professional discourse. *Information Processing & Management, 59*(5), Article 103000. https://doi.org/10.1016/j.ipm.2022.103000
Liu, L., Fang, J., Yang, L., Han, L., Hossin, Md. A., & Wen, C. (2023). The power of talk: Exploring the effects of streamers' linguistic styles on sales performance in B2B livestreaming commerce. *Information Processing & Management, 60*(3), Article 103259. https://doi.org/10.1016/j.ipm.2022.103259
Loughran, T., & Mcdonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance, 69*(4), 1643–1671. https://doi.org/10.1111/jofi.12162

Lu, C., Bu, Y., Wang, J., Ding, Y., Torvik, V., Schnaars, M., et al. (2019). Examining scientific writing styles from the perspective of linguistic complexity. *Journal of the Association for Information Science and Technology, 70*(5), 462–475. https://doi.org/10.1002/asi.24126

Lu, X., Casal, J. E., Liu, Y., Kisselev, O., & Yoon, J. (2021). The relationship between syntactic complexity and rhetorical move-steps in research article introductions: Variation among four social science and engineering disciplines. *Journal of English for Academic Purposes, 52*, Article 101006. https://doi.org/10.1016/j.jeap.2021.101006

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the Advances in Neural Information Processing Systems* (pp. 4765–4774). Long Beach, CA, USA: Curran Associates.

Marino Fages, D. (2020). Write better, publish better. *Scientometrics, 122*(3), 1671–1681. https://doi.org/10.1007/s11192-019-03332-4

Matthews, P., & Glitre, K. (2021). Genre analysis of movies using a topic model of plot summaries. *Journal of the Association for Information Science and Technology, 72*(12), 1511–1527. https://doi.org/10.1002/asi.24525

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 12*(8), 639–646.

Merton, R. K. (1974). *The sociology of science: Theoretical and empirical incvestigations (4. dr.)*. Univ. of Chicago Pr.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*(2), 224–240. https://doi.org/10.1002/sce.10066

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*(6), 615–631. https://doi.org/10.1177/1745691612459058

Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers. *System, 34*(4), 566–585. https://doi.org/10.1016/j.system.2006.08.003

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492–518.

Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. Continuum.

Peng, H., Ke, Q., Budak, C., Romero, D. M., & Ahn, Y. Y. (2021). Neural embeddings of scholarly periodicals reveal complex disciplinary organizations. *Science Advances, 7*(17), eabb9004. https://doi.org/10.1126/sciadv.abb9004

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169–190. https://doi.org/10.1017/S0140525×04000056

Qiu, Y., Ding, S., Tian, D., Zhang, C., & Zhou, D. (2022). Predicting the quality of answers with less bias in online health question answering communities. *Information Processing & Management, 59*(6), Article 103112. https://doi.org/10.1016/j.ipm.2022.1031

Rangel, F., & Rosso, P. (2016). On the impact of emotions on author profiling. *Information Processing & Management, 52*(1), 73–92. https://doi.org/10.1016/j.ipm.2015.06.003

Roland, M. C. (2009). Quality and integrity in scientific writing: Prerequisites for quality in science communication. *Journal of Science Communication, 8*(2), A04. https://doi.org/10.22323/2.08020204

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back propagating errors. *Nature, 323*(6088), 533–536.

Salter, L., & Hearn, A. (1997). *Outside the lines: Issues in interdisciplinary research*. McGill-Queen's Press-MQUP.

Shah, C., Du, Q., & Xu, Y. (2022). Enhanced TabNet: Attentive interpretable tabular learning for hyperspectral image classification. *Remote Sensing, 14*(3), 716.

Shi, D., Liu, W., & Wang, Y. (2023). Has China's Young thousand talents program been successful in recruiting and nurturing top-caliber scientists? *Science (New York, N.Y.), 379*(6627), 62–65. https://doi.org/10.1126/science.abq1218

Stamatatos, E. (2009). A survey of modern authorship attribution methods. *Journal of the American Society for Information Science and Technology, 60*(3), 538–556.

Sidorov, G., Velasquez, F., Stamatatos, E., Gelbukh, A., & Chanona-Hernández, L. (2013). Syntactic Dependency-Based N-grams as Classification Features. In I. Batyrshin, & M. G. Mendoza (Eds.), *Advances in Computational Intelligence* (pp. 1–11). Springer Berlin Heidelberg.

Sletten, E. M., & Bertozzi, C. R. (2009). Bioorthogonal chemistry: Fishing for selectivity in a sea of functionality. *Angewandte Chemie International Edition, 48*(38), 6974–6998.

Smaldino, P. E., Turner, M. A., & Contreras Kallens, P. A. (2019). Open science and modified funding lotteries can impede the natural selection of bad science. *Royal Society Open Science, 6*(7), Article 190194. https://doi.org/10.1098/rsos.190194

Song, N., Chen, K., & Zhao, Y. (2023). Understanding writing styles of scientific papers in the IS-LS domain: Evidence from abstracts over the past three decades. *Journal of Informetrics, 17*(1), Article 101377. https://doi.org/10.1016/j.joi.2023.101377

Sullivan, D. L. (1996). Displaying Disciplinarity. *Written Communication, 13*(2), 221–250. https://doi.org/10.1177/0741088396013002003

Sun, K., Liu, H., & Xiong, W. (2021). The evolutionary pattern of language in scientific writings: A case study of Philosophical Transactions of Royal Society (1665–1869). *Scientometrics, 126*(2), 1695–1724. https://doi.org/10.1007/s11192-020-03816-8

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Tagliacozzo, R. (1978). Some stylistic variations in scientific writing. *Journal of the American Society for Information Science, 29*(3), 136–140. https://doi.org/10.1002/asi.4630290306

Teich, E., Degaetano-Ortlieb, S., Fankhauser, P., Kermes, H., & Lapshinova-Koltunski, E. (2016). The linguistic construal of disciplinarity: A data-mining approach using register features. *Journal of the Association for Information Science and Technology, 67*(7), 1668–1678. https://doi.org/10.1002/asi.23457

Thornborrow, J., & Wareing, S. (1998). *Patterns in language: An introduction to language and literary style*. Routledge.

van den Besselaar, P., & Mom, C. (2022). The effect of writing style on success in grant applications. *Journal of Informetrics, 16*(1), Article 101257. https://doi.org/10.1016/j.joi.2022.101257

Winch, P. (2015). *The idea of a social science and its relation to philosophy*. Routledge. https://doi.org/10.4324/9780203820766 (0 ed.).

Wu, H., Zhang, Z., & Wu, Q. (2021). Exploring syntactic and semantic features for authorship attribution. *Applied Soft Computing, 111*, Article 107815. https://doi.org/10.1016/j.asoc.2021.107815

Xia, E., & Rao, G. (2020). Diachronic Changes and Multi-dimensional Quantitative Analysis of Sentence Complexity in Contemporary Chinese Written Language. In *Chinese Lexical Semantics: 21st Workshop, CLSW 2020, Hong Kong, China, May 28–30, 2020, Revised Selected Papers* (pp. 618–630). Springer-Verlag. https://doi.org/10.1007/978-3-030-81197-6_52.

Yang, S., Zhou, C., & Chen, Y. (2021). Do topic consistency and linguistic style similarity affect online review helpfulness? An elaboration likelihood model perspective. *Information Processing & Management, 58*(3), Article 102521. https://doi.org/10.1016/j.ipm.2021.102521

Yore, L. D., Hand, B. M., & Florence, M. K. (2004). Scientists' views of science, models of writing, and science writing practices. *Journal of Research in Science Teaching, 41*(4), 338–369. https://doi.org/10.1002/tea.20008

Zheng, R., Li, J., Chen, H., & Huang, Z. (2006). A framework for authorship identification of online messages: Writing-style features and classification techniques. *Journal of the American Society for Information Science and Technology, 57*(3), 378–393. https://doi.org/10.1002/asi.20316

Zhu, W. (2004). Faculty views on the importance of writing, the nature of academic writing, and teaching and responding to writing in the disciplines. *Journal of Second Language Writing, 13*(1), 29–48. https://doi.org/10.1016/j.jslw.2004.04.004