English for Specific Purposes 71 (2023) 161–177

Contents lists available at ScienceDirect

English for Specific Purposes

journal homepage: http://ees.elsevier.com/esp/default.asp

Toward an empirical understanding of formality: Triangulating corpus data with teacher perceptions

Tülay Dixon [a,*], Jesse Egbert [b], Tove Larsson [b], Henrik Kaatari [c], Elizabeth Hanks [b]

[a] Oxford College of Emory University, 801 Emory St, Oxford, GA, 30054, USA
[b] English Department, Northern Arizona University, Box 6032, Flagstaff, AZ 86011, USA
[c] University of Gävle, 801 76 Gävle, Sweden

ARTICLE INFO

Article history:

Keywords:
Academic writing
Formality
Informality
Register
Linguistic variation
Situational characteristics

ABSTRACT

Academic writing is often referred to as "formal," but the teaching and assessment of formality can be challenging as formality has been conceptualized in many ways. The goal of this study is to explore the elusive construct of formality in the context of academic writing, especially with regard to what formality means to academic writing instructors. We used instructors' perceptions of formality (i) to identify relationships between the use of linguistic features in academic texts and perceptions of formality and (ii) to determine the extent to which the situational characteristics of texts (e.g., differences in audience, purpose, and discipline) are related to perceptions of formality. Specifically, we asked 72 academic writing instructors to rate the formality level of 60 short academic text excerpts on a five-point scale. The excerpts were sampled from two publication types (university textbooks, journal articles) in three disciplines (psychology, biology, history). Overall, the results indicate that perceptions of formality can be explained by both linguistic features and situational characteristics. As linguistic features and situational characteristics are intertwined, differences in perceptions of formality seem to be functionally motivated. Implications for the teaching of academic writing are discussed.

1. What is formality?

The concept of formality is integral to academic writing, and resources on academic writing often include advice on how to write more formally (Chang & Swales, 1999; Dixon, 2022a; Hyland & Jiang, 2017). Yet, as outlined below, there seems to be little to no consensus about what it means for academic writing to be formal, which makes formality a challenging construct to teach and assess in the academic writing classroom (Liardét, Black, & Bardetta, 2019).

In the following sections, we provide examples of discourse in which academic writing is described as formal to show various conceptualizations of formality and to argue for the need to empirically examine formality. These discussions also cover the findings of the somewhat limited existing empirical research on perceptions of formality. We then list the research questions that guide the current study—questions that can help move toward a more empirically informed understanding of formality.

* Corresponding author.
E-mail address: tulay.orucu.dixon@emory.edu (T. Dixon).

https://doi.org/10.1016/j.esp.2023.04.006
0889-4906/

1.1. Formality referring to various aspects of writing

The term *formality* has been used to describe and refer to different aspects of writing such as its organization, tone, or content. Excerpt (1) below uses formality to refer to the organization of a text, stating that formal writing should have a logically organized structure. Excerpt (2) presents formality as a tone adopted by writers, recommending that academic writing should be formal rather than conversational or casual. Excerpt (3) discusses formality in terms of the content of writing, urging academic writers to provide a synthesis of research rather than their own impressions. These excerpts were chosen among the ones that were surveyed because they succinctly exemplify the varying conceptualizations of formality that are the focus of this study.

(1) "Unlike fiction or journalistic writing, the overall structure of academic writing is formal and logical. It must be cohesive and possess a logically organized flow of ideas; this means that the various parts are connected to form a unified whole" (USC Libraries, n.d., para. 3).
(2) "The tone used in academic writing is usually formal, meaning that it should not sound conversational or casual" (Monash University Library Guide, n. d., para. 4).
(3) "[I]n the case of academic writing, informal language would make the writing sound like the content is only your impression rather than a scholarly synthesis of research, analysis and critical thinking" (Learning Development, n.d.).

As is clear from these excerpts, the construct of formality has different layers as it can refer to the organization, tone, or content of a text. To our knowledge, no empirical research has attempted to untangle these different layers of formality by examining whether the level of formality can be perceived differently when referring to, for example, the content of a text versus the tone of a text.

In addition to the organization, tone, and content of writing, the term *formality* has also been used to refer to the language of writing, which is discussed next.

1.2. Formality as the language of writing

When formality is viewed as the language of writing, it is often conceptualized in terms of the presence or absence of specific linguistic features. Excerpts (4), (5), and (6) below are examples in which various linguistic features are labeled formal or informal. The fact that each excerpt focuses on a different set of linguistic features leads to the natural conclusion that there are varying intuitions and preferences about formality.

(4) "A writer's choices determine whether a document is formal or informal. The use of slang, abbreviations, nonstandard grammar, lots of exclamation points, and a chatty tone are marks of informality. Passive verbs, big words, antiquated expressions, and correct or even stilted grammar signal formality" (The University of Chicago Press, n.d., para. 2).
(5) "You would no doubt be able to sense that of the three verbs *choose*, *select*, and *pick out* (a phrasal verb), *pick out* is the least formal and *select* is most formal" (Kolln & Gray, 2017, p. 175).
(6) "[F]ew readers notice when you violate these elegant options [avoiding split infinitives, using *whom* as the object of a verb or preposition, not ending sentences with prepositions, and using the singular with *none* and *any*], but some do when you observe them, because doing so makes your writing seem just a bit more self-consciously formal" (Williams & Bizup, 2017, p. 16).

Empirical studies on perceptions of formality are few and far between, but the available literature tends to also highlight different linguistic features as formal or informal. Liardét et al. (2019) had three academic writing instructors rate undergraduate student writing as less versus more formal and then examined these texts in terms of their linguistic features. Despite being infrequent, the following features were exclusively found in texts perceived to be less formal: delexical verbs (e.g., *get*, *have*), contractions, personal pronouns, and personalized mental processes (i.e., *hope*, *believe*, *think*). The following features were found in both more formal and less formal texts but were more common in less formal texts: clause-level grammar issues (e.g., subject-verb agreement issues, fragments, and run-ons), word choice errors, "colloquial language" (e.g., *plenty of*, *lots of*, *obnoxious*), "informal descriptors" (i.e., *a lot*, *also*, *very*, *much*), obligation modals (i.e., *should*, *must*, *have to*), abbreviations, and phrasal verbs.

Larsson et al. (forthcoming) examined undergraduate students' perceptions of formality for six linguistic features with an 'academic' and 'non-academic variant'. Of the six features, three of them were associated with proscriptive commentary (non-academic variants are followed by academic variants): split infinitives, non-split infinitives; contracted forms, non-contracted forms; and exclamation points, no exclamation points. The remaining three features were those commonly used in academic writing: adjectival modification, prepositional modification, and nominal modification. The non-academic variant of these three features was a relative clause because clausal embedding is more common in speech than academic writing (Biber, Gray, & Poonpon, 2011). Except for split infinitives, the academic variants were considered to be more formal, but this trend was most clear for contractions and exclamation marks. These results indicate that descriptive features (i.e., adjectival modification, prepositional modification, and nominal modification) do not seem to be as salient to students as those that are often subject to proscriptive commentary.

1.3. Formality as a situational characteristic of a register

Formality has also been conceptualized as a situational characteristic of a register. A register, according to Biber and Conrad (2019), is a situationally defined and culturally recognized variety of language. A register can be described in terms of its

situational characteristics, including its purpose (e.g., to inform, to advertise, to critique), mode of communication (e.g., written, spoken), and target audience (e.g., general public, a specialist audience), among many others.

Registers can also be described for their pervasive lexico-grammatical features. Some lexico-grammatical features occur frequently in a given register because they serve specific discourse functions required by the situational characteristics of the register (Biber & Conrad, 2019). For example, one of the communicative purposes of journal articles is to concisely disseminate highly technical knowledge, which requires the frequent use of nouns as nouns help package technical information in an economical way (Egbert, 2014). In sum, a critical component of a register analysis is the functional interpretation of the linguistic features with respect to the situations in which they are used. This conceptualization of the register framework is illustrated in Figure 1.

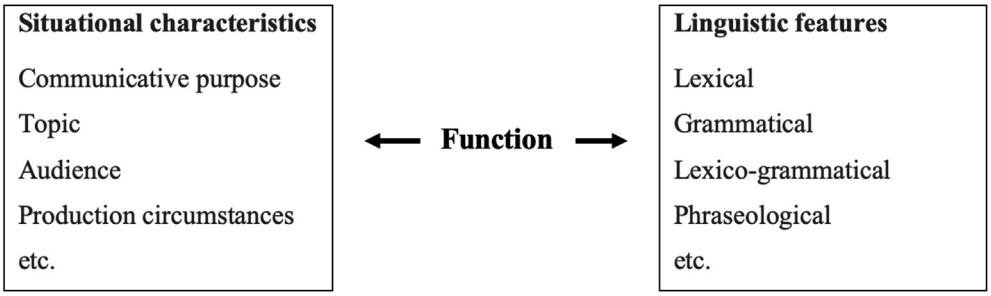


Figure 1. Register analysis framework (adapted from Biber and Conrad, 2019).

Using situational characteristics of registers as the basis of formality, Larsson and Kaatari (2019) conceptualize registers as a cline from more formal to less formal: academic prose, popular science, news, fiction, and conversation (see Figure 2). Based on the situational characteristics, they argue that academic prose, for example, can be considered more formal than popular science, as academic prose is written for a specialist audience whereas popular science is written for a non-specialist audience. Another example comes from Smith (2019), who states that news writing is more formal than personal blogs because these two registers differ in their situational characteristics. While news writing is "highly constrained and associated with media institutions that seek to maintain a professional ethos," personal blogs "are not formally associated with professional institutions and are thus able to discuss a wide variety of topics from serious to frivolous" (Smith, 2019, p. 18). Because of the differences mainly in topic and purpose, Smith (2019) claims that "news writing is almost always more formal than blog writing" (pp. 39–40).

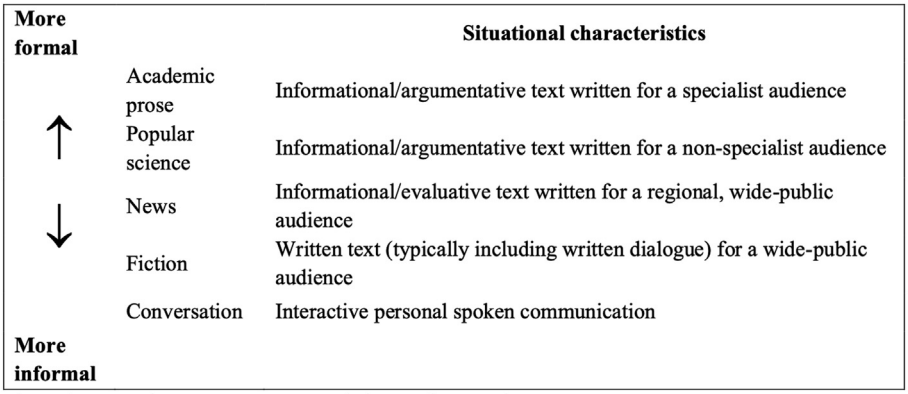


Figure 2. The registers mapped onto the formal-informal cline (adopted from Larsson, 2019; Larsson & Kaatari, 2019).

To our knowledge, Larsson et al. (forthcoming) is the only study that empirically examined the relationship between situational differences and perceptions of formality. The undergraduate students who participated in the study perceived the register of academic articles as more formal than blog posts. Larsson et al. (forthcoming) also examined how situational differences interact with linguistic features to form perceptions of formality. For example, the students perceived the use of an exclamation mark in an academic article as less formal than the use of an exclamation mark in a blog post. They explain this result by referring to *markedness theory*, which refers to the use of a linguistic feature not expected in a given situation. Likely because the use of an exclamation point is less expected in academic writing, students perceived the use of an exclamation point in a blog post as more formal than an exclamation point that appears in academic writing.

Although not a perceptual study, Heylighen and Dewaele's (2002) is yet another study that explored the construct of formality in terms of the situational and linguistic characteristics of texts. To Heylighen and Dewaele, formal writing does not

tolerate ambiguity and fuzzy expressions; thus, it requires context-independent language. Informal writing, which they label *contextuality*, assumes shared background knowledge with the readers (or the interlocutors) and, thus, can use ambiguous and context-dependent language. Based on this understanding of formality, they propose a mathematical formula (see below) that can estimate the level of formality of a given text. The formula computes a value that is between zero and 100; the higher the value is, the more formal the text is.

F = (noun freq. + adjective freq. + preposition freq.+ article freq. − pronoun freq. − verb freq. − adverb freq. − interjection freq. + 100)/2 (Heylighen & Dewaele, 2002, p. 309)

Using the formula, Heylighen and Dewaele calculated formality scores for nine registers that differ in their situational characteristics and ordered them from least formal to most formal: phone conversations, conversations, spontaneous speeches, interviews, imaginative writing, prepared speeches, broadcasts, writing, and informational writing. Although this ordering of registers on a continuum of formality makes intuitive sense, it is unclear whether it would align with *perceptions* of formality. An empirical investigation of perceived formality is one of the goals of the current study.

Thus far, we have discussed some different uses of the term *formality* and the findings of limited research on perceptions of formality. In sum, the challenges we face regarding this term are as follows:

- Empirically, it is unclear whether formality should be conceptualized as a single construct or as a combination of constructs referring to various aspects of writing (e.g., content, language).
- It is unclear which linguistic features are considered formal or informal because different linguistic features have been highlighted as formal or informal.
- It is unclear which situational characteristics make a text more or less formal.

As a result of these challenges, it is difficult to determine what the construct of formality means within the context of academic writing, which, in turn, makes formality a difficult construct to teach and assess reliably.

1.4. The current study

Despite its various conceptualizations, formality seems to be a useful term, as indicated by the ubiquitous references to it in academic contexts. Therefore, the present study seeks to move toward a more empirical understanding of formality in the context of academic writing. We focus on academic writing instructors' perceptions of formality as they are often the first point of contact for students learning to write for academic purposes.

As illustrated in Sections 1.1 and 1.2, formality can refer to different aspects of writing such as its tone, content, and language. Untangling these different layers of formality is one of the aims of the current study. Thus, we first explore whether academic writing instructors perceive the level of formality in a text excerpt differently when formality refers to, for example, the language of the excerpt versus the content of the excerpt.

As discussed in Section 1.2, certain linguistic features have been associated with formality, but there seems to be no consensus as to which linguistic features are perceived as formal or informal. Thus, the current paper examines the perceived formality level of a wide range of linguistic features within the context of academic writing, contributing to the small body of research in this domain.

Finally, formality has also been conceptualized as the situational characteristics of a register (See Section 1.3). Thus, we aim to better understand the extent to which differences in situational characteristics affect perceptions of formality. Academic writing is a broadly defined register category, containing multiple sub-registers that differ in their situational characteristics. Even a narrowly defined academic register such as journal articles varies in its situational and, thus, linguistic nature depending on disciplinary differences (e.g., biology versus history) and research paradigms (e.g., theoretical versus quantitative) (Egbert, 2014; Gray, 2015). Such differences in situational characteristics raise questions such as: Is a history journal article perceived to have the same level of formality as a biology journal article? Is textbook writing considered less formal than journal article writing? The current study aims to answer such questions. Finally, it remains unclear whether perceptions of formality may change depending on whether readers know the register category of the text that they are reading. Therefore, we also examine whether explicit knowledge of the register categories influences writing instructors' perceptions of formality.

The following research questions (RQs) guide the current study:

1. Is formality a single construct or does it refer to a combination of multiple constructs (i.e., language, content, tone)?
2. What linguistic features are associated with writing that is perceived to be formal?
3. To what extent do the situational characteristics of a text affect perceptions of formality?
   (a) To what extent does publication type (i.e., textbooks vs. journal articles) and discipline (i.e., biology, psychology, history) affect perceptions of formality?
   (b) To what extent does knowing the publication type of a text affect perceptions of formality?

2. Method

This section describes the methods of the study, including (i) the survey used to examine academic instructors' perceptions of formality, (ii) the sample of instructors who took the survey, (iii) the 60 text excerpts of about 150 words that were rated by the instructors (i.e., the corpus), and (iv) the statistical analyses used to answer the RQs.

2.1. Survey design and participants

To examine academic writing instructors' perceptions of formality, we asked 72 instructors to complete an online survey. In this survey, the instructors read short academic text excerpts of about 150 words and then rated their level of formality on a five-point scale (0–4). These excerpts were carefully selected to represent two publication types (textbooks and journals articles) in three disciplines (biology, history, psychology), as described in Section 2.2.

The instructors taking the survey were asked to rate the text excerpts on three formality scales: *language*, *tone*, and *content*. This component of the survey allowed us to examine whether formality is a single construct or a combination of constructs referring to various aspects of writing (RQ1). An example task from the survey can be seen in Figure 3. It should be noted that (i) instructors were not asked to rate the formality of text excerpts as it relates to organization because these excerpts were short, not full-length publications; and (ii) the instructors also rated the texts on perceptual scales other than formality (e.g., readability, clarity, interactivity), but we only use the formality ratings to address the research questions that motivate this study.

The text is from a textbook.

Within the larger world that is the biosphere are found many different smaller realms, worlds within worlds as it were. In the sphere of the shadow and darkness lurk the plant-like creatures we know as fungi. Possessing traits of both plants and animals, they are neither. The study of this strange and often bizarre kingdom of creatures is called mycology.

The origins of fungi are unclear, and there is no clear consensus on many of the details of their development. In contrast to plants and animals, the early fossil record of the fungi is meager at best. The body structures of fungi are soft and fleshy and mostly microscopic, not ideal precursors for fossilization. Very few large fungal bodies, such as mushrooms, have ever been found. Fossil fungi are often difficult or impossible to identify and can only be loosely compared to living fungal species.

The language of the text is . . .

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
clear	o	o	o	o	o
readable	o	o	o	o	o
technical	o	o	o	o	o
formal	o	o	o	o	o

The content of the text is . . .

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
abstract	o	o	o	o	o
complex	o	o	o	o	o
informational	o	o	o	o	o
formal	o	o	o	o	o

The tone of the text is . . .

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
unemotional	o	o	o	o	o
serious	o	o	o	o	o
impersonal	o	o	o	o	o
unbiased	o	o	o	o	o
interactive	o	o	o	o	o
formal	o	o	o	o	o

Figure 3. An example task for one of the texts in the survey.

The survey was delivered in two versions. In Survey 1, the instructors were told whether the text excerpt was from a textbook or a journal article, as seen in Figure 3. In Survey 2, the information regarding the publication type was not given to the instructors; that is, the red text in Figure 3 was removed. These two surveys allowed us to examine whether knowing the publication type has an effect on perceptions of formality (i.e., RQ3b). As a result of this design, six different formality ratings emerged from the two surveys (See Table 1).

Table 1
Six formality ratings that emerged from the surveys.

Survey 1 (publication type known)	Survey 2 (publication type unknown)
Language formality	Language formality
Content formality	Content formality
Tone formality	Tone formality

After piloting Surveys 1 and 2 with six academic instructors, we made adjustments to the surveys and delivered it via Qualtrics. Participants were first randomly assigned to Survey 1 or Survey 2. They were then randomly assigned 12 text excerpts to rate (including two training excerpts). To mitigate task-order effect, the excerpts were presented in a random order. Each excerpt was about 150 words and rated by four to seven instructors. The median of the ratings was used to give each text a formality score. Medians were chosen over means because they are less sensitive to outliers and, thus, can more accurately represent the central tendency of ordinal data (Loewen & Plonsky, 2016).

In addition to the perceptual ratings, we collected qualitative data by asking the instructors what characteristics of the text excerpts affected their formality ratings. This qualitative data was collected for six excerpts in the corpus (one from each register category in Table 2) and is presented in the results section with respect to each of the RQs.

Table 2
Number of text excerpts in the corpus.

	History	Psychology	Biology	Totals
Journal Articles	10	10	10	30
Textbooks	10	10	10	30
Totals	20	20	20	60

All instructors who participated in the final survey ($n = 72$) had experience teaching academic writing courses at the university level in the United States. Forty-six of the instructors identified as female, twenty-four as male, and one as non-binary. One participant preferred not to disclose their gender identity. The average instructor had about five years of teaching experience ($M = 5.36$; *Median* $= 3.5$) and was about 41 years old ($M = 41.72$, *Median* $= 40$). Fifty-seven of the instructors spoke English as their first language (or as one of their first languages). Fifteen instructors spoke other languages as their L1, including Bengali, Chinese, German, Kazakh, Korean, Russian, Spanish, Turkish, and Ukrainian.

2.2. Corpus description

Table 2 provides an overview of the 60 text excerpts that were rated by the instructors. These excerpts were about 150 words and extracted from parts of journal articles and textbooks. They represent varying situational characteristics in academic writing, including two publication types (textbooks and journals articles) in three disciplines (biology, history, psychology). Textbooks and journal articles were selected because they differ in their communicative purpose and target audience. Journal articles are written by scholars for other scholars to disseminate novel technical knowledge, whereas textbooks are written by scholars for students and serve as introductory reading materials. In addition to the situational differences in communicative purpose and target audience, disciplinary differences were

also considered. Three disciplines (history, psychology, and biology) were chosen to represent the humanities, social sciences, and sciences. Because the corpus represented varying situational characteristics, we were able to examine the extent to which situational characteristics affect perceptions of formality (i.e., RQ3).

To be able to sample excerpts from full-length publications, we first identified three corpora that included journal articles and textbooks in the targeted disciplines: Academic Written English (AWE), Historical Academic Writing Corpus (HAWC), and Duolingo University Textbooks in English (DUTE), compiled by Egbert (2014), Dixon (2022b), and Egbert and Burch (2023) respectively. From these three corpora, we randomly selected 60 textbooks and journal articles.[1] We then extracted about 150 words from each of the 60 texts; that is, each text contributed *only one* excerpt to the corpus in Table 2. We used different sampling techniques for journal articles and textbooks, which we explain next.

For biology and psychology journal articles, we alternated among the four sections of the articles (i.e., introduction, methods, results, and discussion). For example, we extracted 150 words from the introduction section of the first article, 150 words from the method section of the second article, 150 words from the results section of the third article, and 150 words from the discussion section of the fourth article, repeating this process until all excerpts were collected. For journal articles in history, there were three distinct sections (introduction, main body, and conclusion). The sampling was done by alternating among the three sections. Unlike journal articles, textbooks do not have distinct sections, which is the reason excerpts were sampled from different locations in the books. Accounting for the different sections of the journal articles and textbooks helped minimize *sampling bias*, which refers to the difference between the target domain of texts that a corpus is intended to represent and what texts were actually collected.

To summarize, each publication type and discipline combination is represented by 10 excerpts that are approximately 150 words long ($M = 149$; $Min = 127$; $Max = 173$). Compared to other academic writing corpora, the corpus used in this study is small because each text excerpt needed to be rated by academic writing instructors for perceived level of formality. The short length of the excerpts raises legitimate concerns regarding whether these texts can adequately represent the linguistic characteristics of the targeted population. We address this concern in Appendix B.

2.3. Linguistic analysis of the text excerpts

The second RQ focuses on the linguistic features associated with writing that is perceived to be formal. Before carrying out this analysis, the linguistic profiles of the text excerpts were examined. For this purpose, we used Multi-Dimensional (MD) analysis (Biber, 1988). In MD analysis, the distributions of many linguistic features are analyzed using exploratory factor analysis. Factor analysis is a dimension reduction technique that is used to examine co-occurrence patterns among targeted linguistic features, identifying latent dimensions of language variation across texts and registers. Each dimension is associated with a set of co-occurring linguistic features that serve related functions in texts.

There are two types of MD analyses. In the first type, a full MD analysis is conducted to identify new dimensions with respect to the registers being analyzed (e.g., see Goulart & Wood, 2021, for a list of such MD analyses). In the second type, previously identified dimensions are used to analyze linguistic variation in texts and registers. For example, Staples, Biber, and Reppen (2018) use the dimensions identified by Gardner, Biber, and Nesi (2019) to compare the language that second language writers produce in TOEFL iBT tasks to the language that they produce as part of their academic studies. The latter type is the approach taken in this study, as we used the dimensions identified by Egbert (2014) and calculated the five Egbert (2014) dimension scores for each of the text excerpts in our corpus. These dimensions worked well for capturing the linguistic variation present in our corpus of academic texts as the registers in our corpus are the same as those in Egbert (2014). In the following paragraphs, we give a broad overview of these dimensions, but fine-grained details regarding this original MD analysis can be found in Egbert (2014).

Analyzing the language of three academic publication types (i.e., journal articles, textbooks, and popular science) in two disciplines (biology and history), Egbert (2014) identified five dimensions of linguistic variation, which are listed in Table 3. The first column of the table identifies the label of the dimensions. These labels indicate the discourse functions that are served by the linguistic features associated with the dimension. The second column provides a list of the linguistic features in each dimension. Each linguistic feature is followed by a number in parentheses. These numbers indicate the factor loading of each linguistic feature; that is, how strongly the feature is associated with the dimension. The third column gives examples of academic texts that were found to be typical of the dimensions. The dimensions relevant to perceived formality are explained and exemplified in the results section.

[1] Appendix A lists the journals from which AWE and HAWC sampled articles. The journals were selected with input from experts in the targeted disciplines.

Table 3
An overview of Egbert (2014) dimensions.

Egbert (2014) Dimensions	Linguistic Features	Example academic texts characterized by the dimension
1. Non-technical synthesis vs. Specialized information density	Positive features (Non-technical synthesis): **Non-technical:** COCA Core Vocabulary[2] (1–500) (.61); general adverbs (.59); amplifiers (.43); certainty adverbs (.37); emphatics (.36) **Synthesis:** adverbial conjuncts (.51); phrasal coordinating conjunctions (.39) **Other:** verb *HAVE* (.36); *that*-relative clauses (.36) Negative features (Specialized information density): **Specialized:** technical concrete nouns (−.31) **Information Density:** pre-nominal modifiers (−.73); nouns (−.73); agentless passive voice (−0.42)	+ Popular Academic Books in Biology − Journal Articles in Biology
2. Definition and evaluation of new concepts	Positive features: **Definition:** present tense (.73); verb *BE* (.59); predicative adjectives (.53); concrete nouns (.30) **Evaluation:** non-finite *to*-clauses controlled by stance adjectives (.68); possibility, permission and ability modals (.67); *that*-clauses controlled by attitudinal adjectives (.57); prediction modals (.38) **Concepts:** demonstrative pronouns (.40); pronoun 'it' (.34) **Other:** Academic lexical bundles (.58) Negative features: NONE	+ Textbooks in Biology
3. Author-centered stance	Positive features: **Author-centered:** communication verbs (.58), mental verbs (.58), suasive verbs (.42); human nouns (.41); cognition nouns (.39); 1st person pronouns (.36) **Stance**: *that*-clauses controlled by communication verb (.53); *that*-clause controlled by likelihood verb (.46); *that*-clause controlled by stance noun (.38); *that*-clause controlled by certainty verb (.30) **Other:** infinitives (.53) Negative features: NONE	+ Popular Academic Books in History
4. Colloquial narrative	Positive features (Colloquial narrative): **Colloquial:** common phrasal verbs (.47) **Narrative:** past tense verbs (.59); 3rd person pronouns (.39); aspectual verbs (.44) **Other:** activity verbs (.54), progressive aspect (.37) Negative features (Non-colloquial): **(Non-)colloquial:** Academic Vocabulary List[3] (−.64)	+ Popular Academic Books in History − Journal Articles in Biology
5. Abstract observation and description	Positive features: **Abstract Observation**: nominalizations (.71); word length (0.65); process nouns (.59); other abstract nouns (.47) **Description**: attributive adjectives (.53); topic adjectives (.47) **Other:** Core Vocabulary (501–3000)[4] (.42) Negative features: **Other:** time adverbials (−.32)	+ Journal Articles History − Popular Academic Books Biology

2.4. Statistical analyses

For all analyses, the unit of observation was a text excerpt; that is, all measurements were taken at the level of the excerpts. In total, 11 measurements were taken for each excerpt: five dimension scores from the MD analysis and six formality scores from the perceptual survey (i.e., language, tone, and content formality when the publication type is revealed and not revealed). To reiterate, each formality score was the median of all the perceptual ratings an excerpt received.

RQ1 explores whether *language*, *content*, and *tone formality* represent the same construct. Thus, the relationships among the three formality scores were examined using Pearson's correlations. RQ2 focuses on the linguistic features that are associated with writing perceived to be formal. To answer this RQ, Pearson's correlations were used to examine the relationship between Egbert (2014) dimension scores and the perceptual formality ratings completed by the instructors. In addition to correlations, a multiple regression analysis was conducted using R (R Core Team, 2022) to determine how much of the variance in perceptions of language formality can be explained by the dimension scores.

RQ3 focuses on the extent to which perceptions of formality are affected by (i) publication type, (ii) discipline, and (iii) whether the publication type was revealed to the participants or not. To answer this RQ, we used linear mixed effects models on survey data. All analyses were conducted using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in *R* (R Core Team, 2022). In the model, the fixed factors (i.e., fixed independent variables) were publication type, discipline, and

[2] The percentage of a text composed of words with frequency rankings between 1 and 500 in the Corpus of Contemporary American English.
[3] The percentage of a text composed of words on the Academic Vocabulary List (see Gardner & Davies, 2014).
[4] The percentage of a text composed of words with frequency rankings between 501 and 3,000 in the Corpus of Contemporary American English.

publication type awareness, and text excerpts were included as a random effect. The fixed factors were examined with an ANOVA using Sattherwaite's method. The group differences within each fixed factor were analyzed with Tukey adjusted post-hoc pairwise comparisons using the *emmeans* package (Lenth et al., 2020). To determine the amount of variance explained by the fixed and random factors included in the model, an R^2 value was calculated using the formulas developed by Nakagawa and Schielzeth (2013). The calculations were done using the *r.squaredGLMM* function of the MuMIn package in *R* (Barton, 2009). The output produced two values: (1) *Marginal* R^2 for the variance explained by the fixed factors and (2) *Conditional* R^2 for the variance explained by the fixed *and* the random factors.

3. Results

RQ1 focuses on whether formality is a single construct or a combination of constructs that refer to various aspects of writing. Thus, academic writing instructors were asked to rate the formality level of the *language*, *content*, and *tone* of the text excerpts, and the relationships among the three formality scores were examined using Pearson's correlations. The results are presented in Table 4.

Table 4
Bivariate correlations between formality ratings.

Survey 1 (publication type known)				Survey 2 (publication type unknown)			
	Content	Language	Tone		Content	Language	Tone
content	–	.77	.83	content	–	.78	.82
language		–	.70	language		–	.82
tone			–	Tone			–

The correlations range from .70 to .83, indicating that the three formality scores are strongly associated (according to Cohen's [1998] benchmarks for the interpretation of correlations). However, at the same time, the correlations are not strong enough to suggest that the three formality scores are identical. Thus, we chose not to aggregate the three constructs when analyzing the data to answer RQ2 and RQ3. This decision was also motivated by the qualitative data collected from the instructors. The data suggested that the instructors were able to perceive differences regarding language, tone, and content formality as they commented on them separately and rated them differently:

> "**The tone** does not strike me as formal due to a lack of complex grammatical structures for the most part which is a 'marker' for me of formality and, to tell the truth for this sample, **the topic itself**. At first I thought it would discuss the reporters, but it shifted to an interesting 'who is who' description of attendees" (*P18 ratings for a journal article excerpt in history from Survey 1; language = 3, content = 2, tone = 1).*

> "**The focus on how attendees looked [the content]** seemed like the writer was focusing on rather informal things" (*P14 ratings for a journal article excerpt in history in Survey 2: language = 2, content = 1, tone = 1*).

RQ2 focuses on the linguistic features that are associated with writing perceived to be formal. To answer this RQ, we examined the relationship between Egbert (2014) dimension scores and the perceptual formality ratings completed by the instructors. We focused only on the *language formality* ratings from Survey 2 (i.e., publication type unknown) for two reasons. First, language formality ratings are related to the linguistic features of a text. Second, in Survey 2, the instructors were not given information regarding the publication type of the text excerpts that they rated; that is, they were able to focus on the linguistic features of the text excerpts without their perceptions of formality being affected by the publication type information. This difference between Survey 1 and Survey 2 made Survey 2 more appropriate to answer RQ2.

The correlations between formality ratings and dimension scores can be seen in Table 5. The top two strongest correlations are between language formality and Dimension 1 ($r = -.40$) and Dimension 4 ($r = -.39$). According to Cohen's (1988) benchmarks for interpreting correlations, these correlations can be considered moderate to large, suggesting that co-occurrences of certain linguistic features are related to perceptions of formality. In addition to correlations, we also conducted a multiple regression analysis (MRA) to determine how much of the variance in perceptions of language formality can be explained by the Egbert (2014) dimension scores. Table 6 presents the results of this MRA analysis. All five dimensions could explain about 36% of the variance in the language formality scores, $R^2 = 0.36$, $F(5, 54) = 6.193$, $p < .001$; Adjusted $R^2 = 0.31$. Dimensions 1, 2, and 4 were significant predictors of language formality scores, with Dimension 4 explaining 15% of the variance, Dimension 1 explaining 12% of the variance, and Dimension 2 explaining 8% of the variance, which gives support to the argument that co-occurrences of certain linguistic features are associated with perceptions of formality. As Dimensions 1 and 4 explain most of the variance and are most strongly correlated with perceptions of formality, we will focus on these two dimensions.

Table 5
Correlations between dimension scores and language formality scores.

	Dimension 1	Dimension 2	Dimension 3	Dimension 4	Dimension 5
Language formality	−.40	−.32	−.13	−.39	−.01

Table 6
Multiple regression results for the model containing five predictor variables of language formality scores.

Predictor	β	B	t	p	R^2
Dimension 1	−0.30	−0.04	−2.581	[a].013	0.12
Dimension 2	−0.27	−0.04	−2.331	[a].024	0.08
Dimension 3	0.05	0.01	0.394	.695	0.01
Dimension 4	−0.41	−0.09	−3.578	[a].001	0.15
Dimension 5	−0.18	−0.03	−1.560	.125	0.01

[a] Notes. β = standardized beta. B = unstandardized beta. * = Significant at $p < .05$.

Table 7 includes text excerpts from the corpus that are representative of Dimension 1 (see the supplementary material for more excerpts). The text excerpts that were perceived to be less formal tended to include more "non-technical synthesis" features, and the text excerpts that were perceived to be more formal tended to include more "specialized information density" features.

Table 7
Excerpts exemplifying Dimension 1 'Non-technical synthesis vs. Specialized information density'.

Less formal Non-technical Synthesis ⟷	More formal Specialized Information Density
Selected features: **Bolded = Core Vocabulary (COCA 1-500)** Red: General adverbs Purple: Emphatics Blue: adverbial conjuncts & coordinating conjunctions	Selected features: **Bolded: nouns** Underlined: pre-modifying nouns Blue: Agentless Passive Voice
Another very **important** perceptual **ability that seems to** develop very **early in** infancy, if **it is not** actually present **at** birth**, is the** perception **of** depth. **The** ability **to see the world in** three dimensions **is called** depth perception. **It's a** handy ability because **without it you would have a hard time** judging **how far away** objects **are . . .** (Excerpt from a psychology textbook)	**Methods** for **traction force microscopy** have been previously described (Sabass et al., 2008; Aratyn-Schaus and Gardel, 2010). In brief, **images** of fluorescent **beads** embedded in the PAA **gel** were aligned to compensate for experimental **drift**, and the **bead displacement field** was calculated between **pairs** of **images** by comparing the unstrained **bead images** obtained after the **cell** had been removed . . . (Excerpt from a biology journal article)

Note. This visualization is simplified and does not account for linguistic features that may be in more than one category.

Some of the non-technical synthesis features (i.e., those whose co-occurrence was considered to be less formal) include the proportion of the text excerpt composed of the 500 most common vocabulary from COCA (Davies, 2008), general adverbs (e.g., *actually*), emphatics (e.g., *very*) and adverbial conjuncts and coordinating conjunctions (e.g., *if*, *because*). That the core vocabulary is common suggests a non-technical discourse. In this non-technical discourse, information and ideas are connected and synthesized using adverbial conjuncts and coordinating conjunctions. The presence of adverbs and emphatics add interaction and emotive emphasis to the information being presented.

Some of the specialized information density features (i.e., those whose co-occurrence was considered to be more formal) include nouns, pre-modifying nouns, and agentless passive voice. In general, nouns and pre-modifying nouns allow for the dense packaging of highly technical information, and agentless passive voice allows writers to shift the focus away from them to what they did as part of the study.

In text excerpts with more specialized information density features, the use of passive voice in particular was salient to the participants and was perceived as formal:

> "Because it is social science, **the use of passive voice** makes it more formal" (*P31 ratings for a journal article excerpt in psychology: language = 3, content = 3, tone = 3*).

> "Use of citations, use of Latin abbreviation (i.e.), **"it has been argued and demonstrated" - passive voice** and the words *argue* and *demonstrate*, other word choices that feel formal (e.g., impartiality, agents, allocation)" (*P7 ratings for a journal article excerpt in psychology: language = 4, content = 4, tone = 4*).

The dense use of nouns was also perceived to be more formal, but not all instructors were able to express this perception by referring to nouns. Instead, they shared example sentences that sounded formal to them. For example, two instructors shared the following sentence to exemplify why a biology journal article excerpt sounded formal to them. The sentence has frequent uses of nouns (bolded) that are modified by prepositional phrases (underlined):

Differences among the four **reserve types** were indicated (K–W = 11.390; p = 0.010), perhaps because of the large **difference** between **parks** and extractive **reserves** in the **ratio** of **deforestation** in the outer and inner **buffer**.

One participant shared why this particular sentence was perceived as formal:

> "Precise and unambiguous language (e.g., deforestation), **use of complex noun phrases** (e.g., Differences among the four reserve types)" (*P7 ratings for a journal article excerpt in biology: language = 4, content = 4, tone = 4*).

The second dimension most strongly correlated with perceptions of formality is Dimension 4 'Colloquial Narrative'. The negative correlation between Dimension 4 scores and formality scores indicates that text excerpts that were perceived to be less formal tended to use linguistic features associated with Colloquial Narrative. Table 8 includes text excerpts from the corpus that are representative of this dimension.

Table 8
Excerpts exemplifying Dimension 4 'Colloquial Narrative'.

Less formal Colloquial Narrative	⟷	More formal Non-colloquial
Selected features: Red = Past tense Green: Phrasal verbs Blue: Activity verbs		Blue = 500 most common academic words
The Augustan period was a time of outstanding creativity. The recuperation of peace and stability certainly was conducive … and destruction had been banished; for good reasons, Augustan Rome looked back to the Athens of Pericles. . . . Instead, and as can be readily imagined, what he did and the way he reshaped things engaged the imagination of his contemporaries and stimulated many responses. (Excerpt from a history textbook)		The criteria for including a study in the meta-analysis are listed in Figure 1 according to the PICOS categories (population, intervention, comparison, outcomes, study design; Moher et al., 2015). One requirement was that a study measured some aspect of specificity in future thinking. As described a priori, we included measures that directly assessed one or both key components of episodic specificity (Tulving, 1972). (Excerpt from a psychology journal article)

Note. This visualization is simplified and does not account for linguistic features that may be in more than one category.

Some of the colloquial narrative features (i.e., those whose co-occurrence was perceived as less formal) include past tense verbs, phrasal verbs, and activity verbs. Their co-occurrence creates a narrative style in which past events are recounted in an easy-to-process way. The non-colloquial end of the dimension, however, is marked by the frequent use of words found on the Academic Vocabulary List (Gardner & Davies, 2014).

That the narrative style is considered informal was also mentioned in the qualitative data. In the following participant comments, all parts except for the bolding are the language found in the text excerpts the participants rated.

> "'She wondered: Would a toddler peering over the rim perceive the dangerous drop-off and draw back? Gibson's inspiration for these experiments occurred while she was picnicking on the rim of the Grand Canyon.'- **Reads like a story**" (*P55 ratings for a textbook excerpt in history: language = 2, content = 2, tone = 1*).

> "'Back in their Cornell University laboratory, Gibson and Walk placed 6- to 14-month-old infants on the edge of a safe canyon' is **narrative-based and not very formal.** (*P69 ratings for a textbook excerpt in psychology: language = 3, content = 3, tone = 3*).

In both Dimension 1 and Dimension 4, vocabulary emerged as an important linguistic feature associated with perceptions of formality. In Dimension 1, the use of 500 most common vocabulary words added to perceptions of informality whereas in Dimension 4, use of words on the Academic Vocabulary List added to perceptions of formality. The importance of vocabulary was commonly mentioned in the qualitative data as well:

> **"Complexity of vocabulary,** punctuation and sentence structure" (*P71 ratings for a textbook excerpt in history: language = 3, content = 3, tone = 3*).

> **"[V]ocabulary was a big factor**—longer words (and words that I don't understand) → more formal" (*P5 ratings for a journal article excerpt in history, language = 2, content = 2, tone = 2*).

RQ3 focuses on the extent to which differences in situational characteristics influence perceptions of formality. To answer this question, a linear mixed effects model was used to examine the extent to which perceptions of formality can be explained by the differences in publication type and discipline as well as knowing the publication type (i.e., Survey 1 vs. Survey 2). For these analyses, we once again use only *language formality* scores. Before proceeding with the results, we provide a descriptive overview of the complete dataset in Table 9.

Table 9
Descriptive statistics organized by discipline, publication type, and survey.

	Survey 1 (Publication Type Known)					
	Journal Article Excerpts			Textbook Excerpts		
	N	*M*	*SD*	*n*	*M*	*SD*
Biology	10	3.70	0.48	10	3.30	0.42
Psychology	10	3.35	0.47	10	2.75	0.63
History	10	3.05	0.50	10	2.80	0.48
	Survey 2 (Publication Type Unknown)					
	Journal Article Excerpts			Textbook Excerpts		
	N	*M*	*SD*	*n*	*M*	*SD*
Biology	10	3.50	0.47	10	2.75	0.59
Psychology	10	3.50	0.41	10	2.45	0.64
History	10	2.60	0.46	10	2.60	0.52

The mean scores in Table 9 indicate the following general trends: (i) journal article excerpts are generally perceived as more formal; (ii) writing in biology is generally perceived as more formal than writing in psychology and history, and (iii) when the instructors knew that the text excerpts they were rating came from journal articles and textbooks, they generally rated them as more formal. These descriptive statistics indicate that situational characteristics of writing are also associated with perceptions of formality.

The trends observed in descriptive statistics are supported by the results from the linear mixed effects model. The three independent variables in the model (i.e., publication type, discipline, and knowing the publication type) could account for about 44% of the variance in perceptions of formality. Of the 44% variance, 32% could be explained by the fixed effects of publication type (16%), discipline (12%), and survey (4%). These findings show that in addition to linguistic features as discussed above, situational characteristics of writing can also be associated with perceptions of formality. An additional 12% of the variance was accounted for by the random effect of text excerpt, indicating that there are also features of text excerpts associated with perceptions of formality that were not accounted for by the variables analyzed in this study.

An ANOVA was used to examine the effect of each independent variable (i.e., publication type, discipline, and knowing the publication type) on perceptions of formality. A significant effect was found for all three variables: publication type, $F(1, 56) = 23.210$, $p = <.001$; discipline, $F(2, 56) = 9.082$, $p < .001$; survey, $F(1, 59) = 8.671$, $p = .005$. To better understand the differences between the groups of each fixed independent variable, post hoc pairwise comparisons were run using Tukey HSD tests. The results from these comparisons are presented in Table 10.

Table 10
Post hoc pairwise comparisons.

	Publication Type Contrasts				
	estimate	*df*	*t-ratio*	*p*	*d*
Journal-Textbook	0.508	56	4.818	<.001[a]	1.29
	Discipline Contrasts				
	estimate	*df*	*t-ratio*	*p*	
Biology-History	0.55	56	4.256	.0002[a]	1.14
Biology-Psychology	0.30	56	2.321	.0610	0.62
History-Psychology	−0.25	56	−1.935	.1385	−0.52
	Survey Contrasts				
	estimate	*df*	*t-ratio*	*p*	
Survey 1-Survey 2	0.258	59	2.945	.0046[a]	0.77

Survey 1 = Publication Type Known; Survey 2 = Publication Type Unknown.
[a] Significant at $p < .05$; *estimate* = estimated mean difference.

Regarding publication type, the mean difference shows that journal article excerpts received an average language formality score that was 0.51 points higher than textbook excerpts, a difference that was statistically significant with a large effect size, according to the benchmarks introduced by Cohen (1988) for interpreting effect sizes: values around 0.2 as small, values around 0.5 medium, and values around 0.8 as large. Regarding disciplinary differences, biology excerpts were rated as most formal, followed by psychology and then history, but the only statistically significant difference with a large effect size was between biology and history excerpts. Last, when the instructors knew the publication type of the excerpts that they were reading, excerpts were rated as more formal. The mean difference between Survey 1 and Survey 2 ratings was 0.25 points on the language formality scale, which was a statistically significant difference with a medium-to-large effect size.

We further examined the change in perceptions of language formality from Survey 1 to Survey 2 with respect to publication types (See Table 11). For journal article excerpts, the higher formality scores from Survey 1 compared to Survey 2 was not statistically significant and had a small effect size, $t(29) = 1.3298$, $p = .194$, $d = 0.24$. For textbook excerpts, the increase was statistically significant with a medium effect size, $t(29) = 2.855$, $p = .008$, $d = 0.52$. These findings indicate that explicit knowledge of the register category influenced perceptions of formality, especially for textbooks. Journal article writing is perceived as formal regardless of whether the participants knew the publication type, but for textbooks, this explicit knowledge made a statistically significant difference in perceptions of formality. The participants of Survey 2 could have ranked textbook excerpts as less formal thinking that these excerpts were from, for example, academic blogs.

Table 11
Language formality scores of publication types in Survey 1 and Survey 2.

	n	*Mdn*	*M [95% CI]*	*SD*	*d [95% CI]*	*t*	*p*
Journal Article (Survey 1)	30	3	3.37 [3.17, 3.57]	.54	0.24 [−0.12, 0.60]	1.3298	.194
Journal Article (Survey 2)	30	3	3.20 [2.97, 3.43]	.61			
Textbook (Survey 1)	30	3	2.95 [2.74, 3.16]	.56	0.52 [0.14, 0.90]	2.855	.008*
Textbook (Survey 2)	30	2.75	2.60 [2.38, 2.82]	.58			

Notes. Survey 1 = Publication Type Known; Survey 2 = Publication Type Unknown.
Degrees of freedom was 29 for the paired *t*-tests. *Significant at $p < .05$.

That knowing the publication type affects perceptions of formality was also discussed in the qualitative data. One participant mentioned rating a text excerpt formal based on the publication type:

> "[T]he sources [are] academic articles, course textbooks, **that's why I believe they are formal**" (*P12 ratings for a textbook excerpt in biology in Survey 1: language = 4, content = 4, tone = 4*).

Another participant indicated that they would not expect informal language in journal articles and textbooks. That is, whether a text excerpt is formal can be determined on the basis of its register.

> "I teach composition where I see a lot of informal writing, and none of these showed characteristics of that (**due to the genres chosen, you wouldn't expect them to**)" (*P49 ratings for a journal article excerpt in biology in Survey 1: language = 3, content = 3, tone = 3*).

Overall, it appears that knowing the publication type heightened the participants' awareness about the formality of the text excerpts, giving support to conceptualization of formality as a situational characteristic.

4. Discussion

The motivation for the current study was to better understand what perceptions underlie the construct of formality for academic writing instructors. We first examined whether instructors' perceptions of formality change when they evaluate the content, language, and tone of a text excerpt separately, finding that these three constructs of formality are not identical. Thus, instead of aggregating these three formality scores, we decided to treat them separately and further explored what underlies language formality.

We examined whether the level of perceived language formality of a text excerpt can be predicted by its situational characteristics or use of linguistic features. Our findings indicate that both linguistic features and situational characteristics can explain variation in instructor perceptions of formality. There is a linguistic aspect of formality as evidenced by (i) the variance explained by the Egbert (2014) dimensions and (ii) the qualitative data in which instructors labeled various linguistic features as formal or informal (e.g., passive voice, technical vocabulary, complex noun phrases). There is also a situational dimension to formality as the publication type and discipline of the text excerpts could explain variation in perceptions of formality. Additionally, the same text excerpts were rated as more formal when instructors knew the publication type of the text excerpts, indicating that the publication type sets expectations for the formality level of a text.

In our study design and statistical analyses, we treated situational and linguistic aspects of formality separately. Our findings, however, indicate that linguistic features and situational characteristics work together in determining the level of perceived formality. For example, a higher frequency of information density features (e.g., nouns, pre-nominal nouns, technical nouns, agentless passive voice) was associated with higher formality scores, and journal article excerpts were perceived as more formal than textbook excerpts. Journal article excerpts, however, also contained more information density features than textbook excerpts (See Figure 4 in which lower scores indicate a higher frequency of information density features). This relationship between journal articles and information-density features is not a coincidence as these features allow researchers to concisely package the highly technical information reported in journal articles.

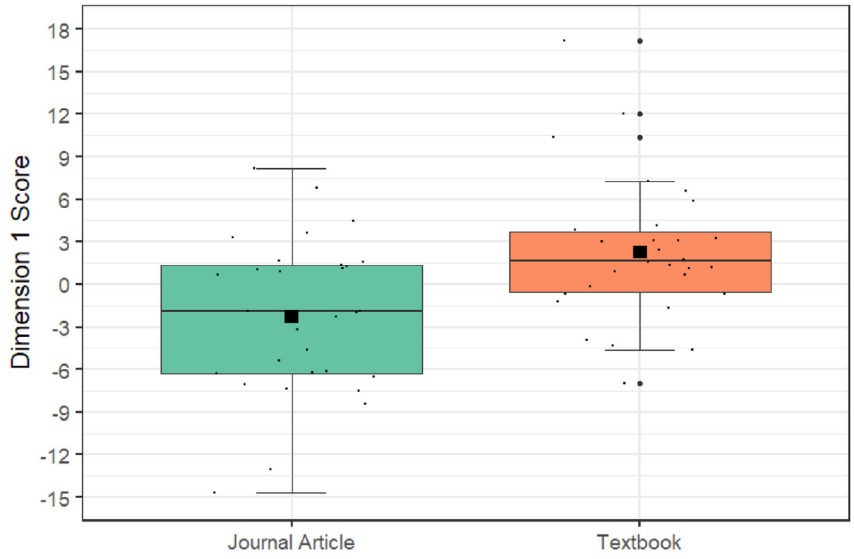


Figure 4. Dimension 1 scores across publication types.
Note. The boxes indicate the middle 50% of the data (25th and 75th percentiles). The solid horizontal lines through the boxes indicate the median scores. The whiskers extending from the boxes show the minimum and maximum scores, excluding the statistical outliers represented by the small black dots.

Another example comes from the functional relationship between narrative features (e.g., past tense verbs, phrasal verbs, and activity verbs) and the situational characteristics of textbooks. A higher frequency of narrative features was associated with lower formality scores, and textbook excerpts were perceived as less formal than journal article excerpts. Textbook excerpts were also the register that tended to contain more narrative features (See Figure 5) likely because such features help create a narrative style in which past events are recounted in an easy-to-process way for students, who are the primary target audience of textbooks. This trend was especially true for biology and psychology for which considerable differences could be observed between textbook and journal article excerpts in terms of the frequency of narrative features (See Figure 5 where higher scores indicate more narrative features.).

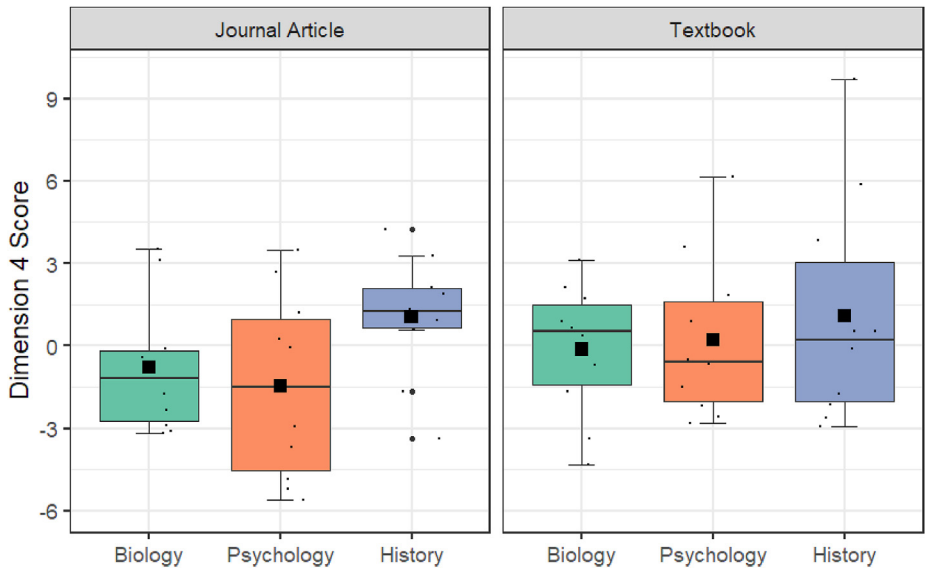


Figure 5. Dimension 4 scores across publication types and disciplines.

Collectively, these findings indicate that although a set of co-occurring linguistic features can be perceived as less formal, they may be the appropriate linguistic features to use in a given register because of the functions that they serve in that register. For example, whereas narrative features were perceived as less formal, they are the linguistic features commonly used in textbooks so that an easy-to-process writing style can be created for students who are being introduced to a scientific

community. Then, labeling certain linguistic features as informal (and, thus, giving the impression that they are features to be avoided in writing) can be misleading as the use of these linguistic features might be required by the situational characteristic of the register in question. Thus, what underlines the construct of formality seems to be the functions served by linguistic features in a given writing situation rather than a binary categorization of linguistic features as formal or informal.

5. Implications, limitations, and directions for future research

Considering the relationships among a text's formality level, linguistic features, and situational characteristics, we recommend that the construct of formality be approached from a register-functional perspective in academic writing classes (see Biber & Conrad, 2019; Biber et al., 2021). That is, we recommend that when offering instruction or feedback, academic writing instructors help students understand the specific functions of linguistic choices within the situational context of their writing, rather than simply encouraging the use of language that is deemed formal. For example, students learning to produce research articles need to be introduced to phrasal and clausal complexity features not because such features are "formal" or "informal" but because they serve different communicative functions in research articles across the disciplines. Whereas phrasal complexity features help deliver information in a more compact manner and are especially appreciated in hard sciences, clausal complexity features "reflect the need for writers outside of the hard sciences to provide more extended explanations of and justifications for arguments and conclusions" (Staples et al., 2016, p. 30). To give another example, if a student's use of first-person pronouns in the discussion section of a research report is perceived as informal, the instructor could go beyond labeling first-person pronouns as informal and follow up with a functional explanation such as, "the use of first-person pronouns is informal *because* in this writing context they place the reader's focus on you and your subjective perspective rather than on the objective findings of your study." In other words, the instructor could emphasize that it is not the mere existence of first-person pronouns that makes a piece of writing informal; it is the function that a first-person pronoun serves in that given writing situation. Helping students understand the functions served by linguistic features can help them develop register awareness, and by extension, an awareness of what it means to write in ways that are situationally appropriate.

The current study is an initial step toward an empirical understanding of the construct of formality within the context of academic writing. As an initial step, it was limited in terms of the number of text excerpts rated by the instructors and the number of instructors who completed the ratings. A larger sample size can help reduce variation, yielding more stable point estimates for the perceptual ratings of formality and the frequency measures of linguistic features. Additionally, the text excerpts were about 150 words. Increasing the length of the excerpts will also help reduce variation.

Although some of the variance in perceptions of formality could be explained by linguistic and situational characteristics of texts, there is variance left unexplained. Thus, the question remains, "What else matters in perceptions of formality?" The current study provided support for the argument that formality is a multifaceted construct that refers to various aspects of writing such as the content, tone, and language of writing. The current study explored language formality in terms of linguistic and situational characteristics, but there is much more to explore regarding tone and content formality. The current study also identified discipline as an important situational characteristic affecting perceptions of formality because biology texts were rated as more formal than psychology and history texts. This finding indicates that technical background knowledge can possibly affect perceptions of formality. In the current study, except for one participant who completed a B.A. in psychology, none of the participants had an educational background in biology, history, or psychology. Thus, future research could examine how perceptions of formality change when texts are rated by experts in a field versus non-experts. Another area future research could examine is whether and to what extent explicit knowledge of phrasal and clausal complexity features affect perceptions of formality since phrasal complexity features are now considered hallmark features of academic writing (see for example Biber et al., 2011; Smirnova, 2022). Although this was not noted as an issue by the pilot participants of the current study, future research could also improve upon the design of the current survey by (i) having different raters evaluate different aspects of formality (i.e., tone, content, language) to ensure that the ratings for one aspect does not influence the ratings for others and (ii) examine the effectiveness of the current wording used in the survey to elicit perceptions of formality. Future studies could also examine perceptions of formality beyond the academic context, contrasting registers that show more pronounced differences in their situational and linguistic characteristics. This line of research will help to better explain the construct of perceived formality as well as the functional relationship between situational context and linguistic characteristics.

Acknowledgments

We are grateful to the editors and anonymous reviewers for their very helpful comments and suggestions. We also would like to thank the academic writing instructors who participated in our study.

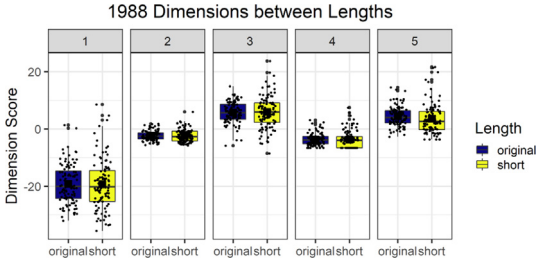
Appendix A. Sampled journals for biology and history articles

Biology	History	Psychology
1. Journal of Natural History	1. American Historical Review	1. American Psychologist
2. Microbial Ecology	2. The Western Historical Quarterly	2. Frontiers in Psychology
3. Journal of Cell Biology	3. Journal of World History	3. Perspectives on Psychological Science
4. American Journal of Physiology	4. Journal of Women's History	4. Psychological Science
5. Conservation Biology		5. Psychological Review

Appendix B. Biber (1988) dimension score differences between short and long texts

Before compiling the corpus described in Section 2.2 and conducting the MD analysis described in Section 2.3, we took additional steps to ensure that 150-word excerpts from textbooks and journal articles are long enough to adequately represent the linguistic characteristics of these targeted registers. We calculated Biber (1988) dimension scores for (i) the original longer texts in the Academic Written English (AWE) corpus and (ii) their shortened versions of about 150 words. We then compared the dimension scores obtained from the longer and shorter texts. The results of this analysis, which can be seen in the table and figure below, indicated more variability in the dimension scores of the shorter texts (i.e., larger standard deviations). However, we found no statistically significant differences between the long and short versions, and the effect sizes were quite small, ranging from 0.02 to 0.19. Thus, we concluded that excerpts of about 150 words were sufficiently long to represent the texts in the MD analysis.

	n	*M* (*SD*)	*CI* around the *M*	*Median*	*SE*	*d* [*CI*]	*p* from *t*-test
Dim 1 Long	100	−19.24(6.9)	[-20.60,-17.88]	−20.09	0.69	0.15[-0.26,0.29]	.9179
Dim 1 Short	100	−19.35(8.83)	[-21.10,-17.60]	−20.26	0.88		
Dim 2 Long	100	−2.38(1.69)	[-2.72,-2.05]	−2.59	0.17	−0.03[-0.31,0.24]	.8372
Dim 2 Short	100	−2.32(2.3)	[-2.78,-1.87]	−2.70	0.23		
Dim 3 Long	100	5.7(3.69)	[4.97,6.43]	5.56	0.37	−0.03[-0.31,0.25]	.8122
Dim 3 Short	100	5.86(5.75)	[4.72,7.00]	5.88	0.58		
Dim 4 Long	100	−3.69(2.14)	[-4.12,-3.27]	−3.88	0.21	0.02[-0.26,0.29]	.9151
Dim 4 Short	100	−3.73(3.16)	[-4.36,-3.10]	−3.96	0.32		
Dim 5 Long	100	4.39(3.54)	[3.69,5.09]	3.96	0.35	0.19[-0.09,0.47]	.1783
Dim 5 Short	100	3.53(5.33)	[2.47,4.58]	2.63	0.53		



Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.esp.2023.04.006.

Data availability

Data will be made available on request.

References

Barton, K. (2009). *Mu-MIn: Multi-model inference*. R Package Version 0.12.2/r18 http://R-Forge.R-project.org/projects/mumin/.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01.
Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.
Biber, D., & Conrad, S. (2019). *Register, genre, and style* (2nd ed.). Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35. https://doi.org/10.5054/tq.2011.244483.
Biber, D., Gray, B., Staples, S., & Egbert, J. (2021). *The Register-Functional Approach to grammatical complexity: Theoretical foundation, descriptive research findings, applications*. Routledge.
Chang, Y.-Y., & Swales, J. M. (1999). Informal elements in English academic writing: Threats or opportunities for non-native speakers? In C. N. Candlin, & K. Hyland (Eds.), *Writing: Texts, processes, and practices* (pp. 145-167).
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
Davies, M. (2008). *The corpus of contemporary American English*. https://www.english-corpora.org/coca/.
Dixon, T. (2022a). Proscribed informality features in published research: A corpus analysis. *English for Specific Purposes, 65*, 63-78. https://doi.org/10.1016/j.esp.2021.09.003.
Dixon, T. (2022b). *Rules of academic writing: A synchronic and diachronic corpus analysis across the disciplines.* [Doctoral dissertation, Northern Arizona University]. *ProQuest Dissertations & Theses Global*.
Egbert, J. (2014) [Doctoral dissertation, Northern Arizona University]. *Reader perceptions of linguistic variation in published academic writing*. ProQuest Dissertations & Theses Global.
Egbert, J., & Burch, B. (2023). Which words matter most? Operationalizing lexical prevalence for rank-ordered word lists. *Applied Linguistics, 44*(1), 103-126.
Gardner, S., Biber, D., & Nesi, H. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics, 40*(4), 646-674. https://doi.org/10.1093/applin/amy005.
Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*(3), 305-327. https://doi.org/10.1093/applin/amt015.
Goulart, L., & Wood, M. (2021). Methodological synthesis of research using multi-dimensional analysis. *Journal of Research Design and Statistics in Linguistics and Communication Science, 6*(2), 107-137. https://doi.org/10.1558/jrds.18454.
Gray, B. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. John Benjamins.
Heylighen, F., & Dewaele, J. M. (2002). Variation in the contextuality of language: An empirical measure. *Foundations of Science, 7*(3), 293-340.
Hyland, K., & Jiang, F. (2017). Is academic writing becoming more informal? *English for Specific Purposes, 45*, 40-51. https://doi.org/10.1016/j.esp.2016.09.001.
Kolln, M., & Gray, L. (2017). *Rhetorical grammar: Grammatical choices, rhetorical effects* (8th ed.). Pearson.
Larsson, T. (2019). Grammatical stance marking in student and expert production: Revisiting the informal-formal dichotomy. *Register Studies, 1*(2), 243-268. https://doi.org/10.1075/rs.18009.lar.
Larsson, T., & Kaatari, H. (2019). Extraposition in learner and expert writing. *International Journal of Learner Corpus Research, 5*(1), 33-62. https://doi.org/10.1075/ijlcr.17014.lar.
Lenth, R. V., Buerkner, P., Herve, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2020). Emmeans (version 1.5.3): Estimated Marginal Means, aka Least-Squares Means. *R package*. https://CRAN.R-project.org/package=emmeans.
Learning Development (n.d.) Formal writing. https://documents.uow.edu.au/content/groups/public/@web/@stsv/@ld/documents/doc/uow195602.pdf.
Liardét, C. L., Black, S., & Bardetta, V. S. (2019). Defining formality: Adapting to the abstract demands of academic discourse. *Journal of English for Academic Purposes, 38*, 146-158. https://doi.org/10.1016/j.jeap.2019.02.007.
Loewen, S., & Plonsky, L. (2016). *An A-Z of applied linguistics research methods*. Palgrave.
Monash University Library Guide (n.d.). Using academic language. Retrieved from https://www.monash.edu/rlo/research-writing-assignments/writing/features-of-academic-writing/academic-language.
Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*, 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x.
R Core Team. (2022). *R: A language and environment for statistical computing [Computer software]*. https://www.R-project.org/.
Smirnova, E. (2022). Clausal complexity of expert and student writing: A corpus-based analysis of papers in social sciences. *Language Learning in Higher Education, 12*(2), 453-475. https://doi.org/10.1515/cercles-2022-2052.
Smith, T. J. (2019). *A comparison of prescriptive usage problems in formal and informal written English*. Doctoral dissertation. Iowa State University. Graduate Theses and Dissertations https://lib.dr.iastate.edu/etd/17566.
Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication, 33*(2), 149-183.
Staples, S., Biber, D., & Reppen, R. (2018). Using corpus-based register analysis to explore the validity of high stakes language exams: A register comparison of TOEFL iBT and disciplinary writing tasks. *Modern Language Journal, 102*(2), 310-332. https://doi.org/10.1111/modl.12465.
The University of Chicago Press. (n.d.). Manuscript preparation, copyediting, and proofreading. https://www.chicagomanualofstyle.org/qanda/data/faq/topics/ManuscriptPreparation/faq0211.html.
USC Libraries (n.d.). Organizing your social sciences research paper. https://libguides.usc.edu/writingguide/academicwriting.
Williams, J. M., & Bizup, J. (2017). *Style: Lessons in clarity and grace* (11th ed.). Pearson.

Tülay Dixon (PhD, Northern Arizona University) is a Visiting Assistant Professor at Oxford College of Emory University where she teaches linguistics and introductory statistics courses. Her research interests include corpus-based analyses of academic discourse, language attitudes, the use of technology in second language education, and research methodology.

Jesse Egbert is Associate Professor of Applied Linguistics at Northern Arizona University where he specializes in register variation and corpus linguistics. He is a General Editor of *Register Studies*, and his recent publications include books on corpus linguistics methods (Cambridge 2020) and corpus design and representativeness (Cambridge 2022).

Tove Larsson is Assistant Professor of Applied Linguistics at Northern Arizona University. She is also affiliated with Uppsala University in Sweden. She specializes in learner corpus research, L2 writing, linguistic complexity, and research methods. She is also currently leading a project on research ethics in quantitative research in the humanities.

Henrik Kaatari received his PhD from Uppsala University in Sweden. His research interests include syntactic/grammatical variation, complexity and learner corpus research. He is currently senior lecturer in English linguistics at the University of Gävle in Sweden.

Elizabeth Hanks is a doctoral student in applied linguistics at Northern Arizona University. Her research uses corpus methods to investigate questions related to developmental writing, vocabulary acquisition, and pragmatics.