

RESEARCH ARTICLE



Style, content, and the success of ideas

Reihane Boghrati¹ | Jonah Berger² | Grant Packard³

¹Arizona State University, Tempe, Arizona, USA

²University of Pennsylvania, Philadelphia, Pennsylvania, USA

³York University, Toronto, Ontario, Canada

Correspondence

Jonah Berger, University of Pennsylvania, Philadelphia, PA, USA.

Email: jberger@wharton.upenn.edu

Abstract

From marketers and consumers to leaders and health officials, everyone wants to increase their communications' impact. But why are some communications more impactful? While some argue that content drives success, we suggest that style, or the way ideas are presented, plays an important role. To test style's importance, we examine it in a context where content should be paramount: academic research. While scientists often see writing as a disinterested way to communicate unobstructed truth, a multi-method investigation indicates that writing style shapes impact. To separate content from style, we focus on a unique class of words linked to style (i.e., function words such as “and,” “the,” and “on”) that are devoid of content. Natural language processing of almost 30,000 articles from a range of disciplines finds that function words explain 4–11% of overall variance explained and 11–27% of language content's impact on citations. Additional analyses examine particular style features that may shape success, and why, highlighting the role of writing simplicity, personal voice, and temporal perspective. Experiments further indicate the causal impact of style. The results suggest ways to boost communication's impact and highlight the value of natural language processing for understanding the success of ideas.

KEYWORDS

automated textual analysis, language, linguistic style, natural language processing, success of ideas

INTRODUCTION

Everyone wants what they say to have impact. Marketers want their advertisements to resonate, politicians want their policies to be adopted, and health officials want their messages to spread. Salespeople want consumers to listen, writers want their stories to get attention, and academics want their work to be cited.

But why do some things succeed in the marketplace of ideas?

We suggest that even above and beyond an idea's content, how it is presented (i.e., presentation style) also plays an important role. To test this possibility, we examine a class of language features that allows us to distinguish between style and content. Along the way, we shed light on specific stylistic choices that are linked to success, and why.

SUCCESS IN THE MARKETPLACE OF IDEAS

There are two major explanations for success in the marketplace of ideas. The first centers on content. Certain things succeed because they are better or more novel than what came previously (e.g., relative advantage; Rogers, 2010). High-speed Internet replaced dial-up because it was faster and easier to use. Einstein's theory of general relativity replaced Newton's law of universal gravitation because it better explained the experimental evidence.

Another possibility, however, has less to do with content and more to do with style. Some have theorized that the manner, or style, with which ideas are presented shapes their impact (Lakoff & Johnson, 1999; Pinker, 2014). But while this seems plausible, little

empirical work has tested this possibility, in part because of challenges in separating content and style.

To test style's importance, we examine it in a context where one might imagine content is paramount and style is not: academic research. Science prides itself on being an objective exercise, where writing is merely a disinterested way to communicate unobstructed truth (American Psychological Association, 2020; Pinker, 2014). The notion is that some discoveries (e.g., general relativity or prospect theory) are simply more novel or groundbreaking than others, and citations are seen as an unbiased measure of such quality (Hamermesh et al., 1982; Hamermesh & Schmidt, 2003; Smart & Waldfogel, 1996). Consequently, if presentation style shapes success even in such a content-focused domain as academic research, it highlights its potential importance across domains.

Testing style's impact, however, is challenging. It is one thing to theorize that certain presentation styles make ideas more successful. But actually, measuring adherence to stylistic approaches, and linking them to a consequential outcome, is difficult.

Further, it can be difficult to separate style from content. Even if papers that write in certain ways (e.g., use more emotional language) are cited more, this could be driven by the subject matter. Papers studying certain topics (e.g., emotions) likely use more language related to those topics (i.e., emotional language), and thus, content, rather than writing style itself, could be driving impact.

FUNCTION WORDS AND LINGUISTIC STYLE

To address these challenges, we focus on a small class of words that play a unique role in language. Function words make up only a tiny portion of the English vocabulary (i.e., ~0.04%; Baayen et al., 1995) but appear in every sentence. They include articles (e.g., *a* and *the*), auxiliary verbs (e.g., *can* and *must*), common adverbs (e.g., *actually* and *basically*), conjunctions (e.g., *so* and *but*), impersonal pronouns (e.g., *it* and *those*), negations (e.g., *cannot* and *have not*), personal pronouns (e.g., *I* and *you*), prepositions (e.g., *on* and *in*), and quantifiers (e.g., *any* and *lots*).

Function words convey little semantic value by themselves, but bind and enrich the nouns, adjectives, verbs, and some adverbs that make up communication (Ireland & Pennebaker, 2010). The word “*the*,” for example, could appear when people are writing about cars (e.g., *the car*), sports (e.g., *the ball*), or a nearly infinite number of other topics. Further, because they are largely meaningless in isolation, function words are often treated as junk in text analysis, tossed out like garbage before analysis is performed (see Berry-Blunt et al., 2021 for a review).

But while they may seem meaningless in some ways, “the ways people use function words reflects their

linguistic style” (Chung & Pennebaker, 2007, p. 247). Consider two people reviewing a restaurant. One person might tend to use sentences like “I enjoyed Jasper Grille” while a second person might tend to write things like “The Jasper Grille was really enjoyable.” While the content of what these two people are saying is largely the same (i.e., they enjoyed the restaurant), they express that same content using different styles. The choice of different function words (i.e., a personal pronoun versus an article and common adverb) gives stylistic nuance to the content delivered.

Indeed, decades of research in linguistics, psychology, and other disciplines refer to function words as “style words” because they reflect aspects of a communicator's linguistic style rather than what is being discussed (Ireland et al., 2011; Lawson & Matz, 2022; Pennebaker, 2011, 2017; Tausczik & Pennebaker, 2010). That is, when considered at document level (e.g., an article, chapter, or conversation), content words indicate *what* people are communicating while function words reflect *how* they communicate (Tausczik & Pennebaker, 2010).

TESTING STYLE'S IMPACT

We suggest that if function words help explain writing's impact, above and beyond the impact of content, it suggests that style shapes the success of ideas. Our theorizing builds on notions of cultural selection (Akpınar & Berger, 2015; Heath et al., 2001; Norenzayan et al., 2006; Schaller & Crandall, 2003). Most research in consumer psychology focuses on individuals, examining how people make choices or react to particular stimuli. But while this work sheds light on important psychological processes, it usually has less to say about why some cultural items (e.g., books, movies, or academic articles) are more successful than others.

Work on cultural selection tries to bridge this gap. Rather than using individuals as the unit of analysis, it examines cultural items themselves, trying to understand success based on how items fit with people. Successful folktales, for example, tend to be minimally counterintuitive in part because minimally counterintuitive stories are more memorable (Norenzayan et al., 2006). Similarly, high arousal content is more likely to go viral (Berger & Milkman, 2012) in part because arousal increases sharing (Berger, 2011). By linking psychological process at the individual level, with the success of cultural items at the collective level, such work tries to demonstrate how individual-level process shape collective outcomes.

Along these lines, as we discuss later, we suggest that certain ways of writing may make academic articles easier to read or remember or make their findings seem more important, novel, or replicable. Similarly, certain words may make methodological choices seem more objective, or findings more conclusive or generalizable. These features, in turn, should impact the likelihood that readers

cite an article and thus shape the number of citations an article receives.

Note, we do not mean to suggest that all function words are the same, or that they all have similar effects. Different types of function words serve different roles. Prepositions and articles are typically used in highly complex professional writing, for example, while pronouns often appear in more informal writing (Biber et al., 2011; Pennebaker et al., 2003; Pennebaker & King, 1999). That said, given function words capture writing style, we start by using them to test the relationship between style and the success of ideas. Then, once this fundamental relationship has been tested, we examine why particular *types* of function words may be linked to citations.

OVERVIEW OF STUDIES

To test our theorizing, we examine tens of thousands of academic articles from a range of disciplines. First, we test our suggestion that style shapes the success of ideas (Study 1). We examine whether writing style (i.e., the function words used) is associated with the number of citations received, even after controlling for an article's content features (e.g., topical focus or research area. Note that we are not suggesting that using *more* function words is necessarily better. Rather, consistent with prior work, we simply suggest that function words reflect linguistic style, and this, in turn, may shape the impact ideas achieve. Given that factors beyond the article text also likely drive citations (e.g., journal, publication year, and author prominence), we test whether the observed effects persist even including a range of controls. To allow for non-linear relationships or interactions between variables, we also examine whether the results hold using a machine learning model optimized for prediction.

Second, we explore three specific ways writing style might shape success. We start by examining whether particular ways of using function words (e.g., certain types of pronouns) are associated with greater citations (Study 2). Then, to test the causal impact, we conduct simple follow-up experiments (Study 3 and 4). We manipulate function words, measure their impact, and test the hypothesized processes.

STUDY 1: STYLE, CONTENT, AND THE SUCCESS OF IDEAS

To begin to explore our suggestion that style shapes the success of ideas, we analyze almost 29,000 academic articles. We use natural language processing to extract both content and style features, and test whether controlling for a variety of content features, and other

factors, style is linked to the prominence research achieves.

Method

We compiled a corpus of full-text peer-reviewed articles from 1990–2018 from five social science disciplines (i.e., psychology, economics, political science, anthropology, and sociology). Online rankings and discipline experts were used to identify the top journals (e.g., *Econometrica*, *Psychological Science*, and *American Journal of Sociology*) in each discipline. Then, we acquired full-text article data, as well as meta-information (e.g., title, issue, and authors), from as many of these journals as possible from JSTOR, a digital library of academic journals (see Table S1 for more detail).

Given the focus on research, we removed non-research articles (e.g., letters to the editor) and articles with missing text. To remove non-research articles, we searched the meta-information for articles whose subjects included words such as “reports,” “comment reply,” or “note” and searched for phrases like “letter to the editor,” “front matter,” or “in this issue” in article titles. In case this missed some non-research articles, we also removed articles with extremely small word counts, calculating mean word count per journal and removing articles three standard deviations below the mean. JSTOR automatically exports PDFs to text files, and some articles were transferred incorrectly, so we removed any article whose word count per page fell below 200 words per page. This left a dataset of 28,774 articles.

Given the focus on writing style, rather than the titles of articles referenced, references were removed. We identified the last occurrence of the word “references” and calculated the concentration of four-digit numbers beginning with “20” and “19” (as calendar year references to the current or prior century). If those numbers occurred more than a reasonable portion of the time (i.e., 0.5%, or one out of every 40 words, which is approximately the number of words in each reference), we removed whatever came after.

To capture writing style, following prior work (Chung & Pennebaker, 2007; Ireland & Pennebaker, 2010) we measured the incidence rate (i.e., proportion of words) of each of the nine function word categories (i.e., conjunctions, prepositions, quantifiers, negations, grammatical articles, personal pronouns, impersonal pronouns, auxiliary verbs, and common adverbs) in each article using Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2015, see Table S2 for descriptive statistics).

To measure the dependent variable, we collected the number of citations each article received from Google Scholar using the Publish or Perish program (Harzing, 2007).

Controls

Citations are obviously driven by many aspects beyond an article's text, so to test alternative explanations and robustness, we control for a variety of other factors.

Various non-language features might impact citations, so we start there. Articles in certain journals may receive more citations, so we included dummy variables to control for journal. Note that this also controls for any general differences in citation rates across disciplines. Older articles should have more opportunity to accrue citations, so we control for publication year. We also control for a number of features shown in prior work to relate to citations, including article length (Stremersch et al., 2007), abstract length and title length (Van Wesel et al., 2014), article order in the journal (Stremersch et al., 2007), number of authors (Chen, 2012), and number of references (Chen, 2012).

To control for author's gender, we relied on methods from prior research (Fox et al., 2016; Topaz & Sen, 2016). Genderize.io uses public census data to map each author's name to "male," "female," or "none." This combined with the `gender_guesser` python package (version 0.4.0) and the national name list (Kaggle, 2017) successfully extracted authors' gender for most of the articles. For the remaining authors, research assistants manually identified whether an author was female or male by searching for them on the internet. We control for the ratio of female authors.

Certain types of articles (e.g., theory pieces) may be cited more, so we use supervised machine learning to classify articles as theory pieces or not. Four research assistants manually coded 700 articles based on whether they included any experiments, data analysis, or an analytical model. Pieces that did not do any of these were classified as theory pieces. This served as a training set. Next, the feature vector for the machine learning approach was generated, including the presence of equations (calculated by the percentage of equal signs in an article), empirical topics (calculated using LDA to identify 17 empirical topics and summing up the topic probabilities for each article), and term frequency-inverse document frequency features (tf-idf) based on articles' main text. Support Vector Machines (SVM; Cristianini et al., 2000) were then applied on the training dataset, achieving 91% accuracy using 10-fold cross-validation. Finally, the trained model identified the article label (i.e., theory or not) for the rest of the articles.

Article content likely impacts citations, so we control for that as well. Certain research topics or areas (e.g., trade policy or self and identity) might be cited more, so we proxy for this using a well-adopted topic modeling method, Latent Dirichlet Allocation (LDA; Blei et al., 2003). Rather than assuming each article covers only one theme, topic modeling allows each article to be represented as a mixture of different topics (e.g., 10% topic A and 7% topic B). Based on coherence scores, an

80-topic approach was used (see Table S9 for topic examples). We included each article's topic probabilities as controls in the model. Consistent with the notion that function words are capturing style and not content, their average absolute correlation with topics is only 0.05.

Main analytic approach

Given citations are counts, and over-dispersed, we used a negative binomial regression to analyze the link between style features and citation count. Because different models use different numbers of predictors, we use sum-of-squared-errors-based adjusted pseudo R^2 for model comparison (called *adj. R^2* hereafter), which adjusts variance explained for the number of predictors in the model. Results are identical if out-of-sample cross-validation is used for model comparison instead. For this robustness check, we use 10-fold cross-validation, and in each iteration, train a model with 90% of the data (training set), apply it against the remaining 10% of the data (test set), and calculate the predicted R^2 on the test set.

Results

Results indicate that above and beyond the variance explained by non-language and content features (model 1, *adj. R^2* = 0.208), adding style features helps explain how many citations articles receive (model 2, *adj. R^2* = 0.217, $F = 2.665$, $p = 0.004$, Table 1; for full details on coefficients and correlations, see Tables S3 and S4). This suggests that function words, which capture linguistic style, explain 4% of the overall variance explained and 11% of language's impact on citations.

These results are intriguing, but one could wonder whether they might somehow be driven by the modeling approach used. While results are the same using a penalized LASSO regression (Tibshirani, 1996) to identify and remove predictors that may introduce collinearity (*adj. R^2* = 0.217 vs. 0.209, $F = 2.658$, $p = 0.004$), maybe there are non-linear relationships between the non-language or content controls and citations, or interactions between these variables that, once included, would wipe out any effect of style features.

To test this possibility, we used machine learning (i.e., a two-layer feed-forward neural network; for example see Hastie et al., 2009). While linear regression or SVM explores linear combination of variables, a two-layer feed-forward neural network is useful because it tries all combinations of variables, including nonlinear ones, to achieve the best prediction. We implemented the neural network using Keras Python package (version 2.4.3), with two hidden layers with 64 and 32 nodes and TanH as the activation function. Given that the output is count data, the output layer's activation function was set to exponential. As the data come from the Poisson distribution, a

TABLE 1 Style words and citations.

	Negative binomial		Machine learning	
	(1)	(2)	(3)	(4)
	Baseline	+Style	Baseline	+Style
Adjusted R^2	0.208	0.217**	0.209	0.234***
Style Features		Yes		Yes
Content Controls				
LDA Topics	Yes	Yes	Yes	Yes
Non-Language Controls				
Publication Year	Yes	Yes	Yes	Yes
Journal	Yes	Yes	Yes	Yes
Article Length	Yes	Yes	Yes	Yes
Abstract Length	Yes	Yes	Yes	Yes
Title Length	Yes	Yes	Yes	Yes
Article Order	Yes	Yes	Yes	Yes
Num Authors	Yes	Yes	Yes	Yes
Author's Gender	Yes	Yes	Yes	Yes
Num References	Yes	Yes	Yes	Yes
Article Type	Yes	Yes	Yes	Yes
Observations	28,774	28,774	28,774	28,774

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ values compared adding style features in each of the three model comparisons. Dummy variable was used for journal with American Economic Review as the base variable. Because LDA topic scores add up to 1.00 for every article, one of the topics was dropped for model considerations.

Poisson loss function was used. Even using this more sophisticated approach, however, including style features still adds additional predictive power (model 3 vs. model 4, *predicted* $R^2 = 0.234$ vs. 0.209, $F = 7.947$, $p < 0.001$).

Overall, these results suggest that function words, which capture linguistic style, explain 4–11% of the overall variance explained and 11–27% of language's impact on citations.

Robustness checks

To further test alternative explanations, we conducted a number of robustness checks. In each, models are compared to the model with all the non-language and content features (i.e., model 1 in Table 1).

Other content features

First, to further control for content features, we included the prevalence of the psychological process dictionaries from Linguistic Inquiry and Word Count (i.e., affective processes, social processes, cognitive processes, perceptual processes, biological processes, drives, time orientations, relativity, personal concerns, and informal language; Pennebaker et al., 2015).

Results remain the same: including style features still adds predictive power (*adj.* $R^2 = 0.220$ vs. 0.211, $F = 2.382$,

$p = 0.011$). Results also remain the same controlling for standard measures of readability (e.g., Flesch reading ease, *adj.* $R^2 = 0.217$ vs. 0.208, $F = 2.794$, $p = 0.003$).

Where authors are from

Second, one could wonder whether the results are driven by where authors are from. Scholars outside the United States may write differently or be cited less (Stremersch & Verhoef, 2005). To test whether this can explain the results, we collected data on author institutions from Elsevier. We identified which institution each author was associated with, and whether the institution was based in the United States. For each academic article, we then calculated the percentage of authors that were United States based and used that as a control.

While this measure was only available for a portion of the papers in the dataset, even controlling for it, including style features still adds predictive power (*adj.* $R^2 = 0.238$ vs. 0.231, $F = 1.885$, $p = 0.049$).

Author prominence

Third, one might wonder whether the results are driven by author prominence. Maybe more prominent authors use particular writing styles, and it is their prominence, rather than the writing style itself, that impacts citations.

This does not seem to be the case. To measure author prominence, we first scraped authors' Google Scholar profiles to retrieve their total citation count. This was only available for 57% of authors in the dataset, so to retain as many articles as possible, we used maximum citation count among all authors of a given article as the measure of prominence. Even including author prominence in the model, including style features still adds predictive power (*adj. R*² = 0.168 vs. 0.100, *F* = 15.440, *p* < 0.001).

We also measured prominence through authors' academic affiliations. We matched author institution data with global institution rankings from [timeshighereducation.com](https://www.timeshighereducation.com) and averaged school rankings over the past 10 years. For authors whose school rank did not appear on the list, the school rank was set to the lowest rank plus one (i.e., 201). Again, author affiliation was only available for a portion of authors in the dataset, so to retain as many articles as possible we used the highest-ranking school from the authors of a given paper as a measure of prominence. Results remain the same (*adj. R*² = 0.228 vs. 0.221, *F* = 2.249, *p* = 0.017).

Overall then, while author prominence certainly helps explain citations, and prominent authors may write differently, function words still explain citations even controlling for these aspects. Further, even controlling for all the robustness check factors in one model, style features still add predictive power (*adj. R*² = 0.187 vs. 0.146, *F* = 4.829, *p* < 0.001; [Table S5](#)).

Discussion

Taken together, results of Study 1 are consistent with the notion that style shapes the success of ideas. Incorporating style features (i.e., how authors use different types of function words) increases the variance explained by 1.0–2.5%, which is 4–11% of the overall variance explained and 11–27% of the variance explained by topical content (i.e., LDA topics). This result suggests that fewer than 500 style words that contain no ideas or content on their own hold up against over 100,000 content words in explaining an idea's success.

Ancillary analysis

One might wonder whether abstracts alone are enough to explain the effects. To test this possibility, we extracted the abstract from each article. The abstracts' texts are available in meta-information files, but they do not always appear the same way in the articles' text (e.g., different spellings are used, or they are mixed up with author information). Consequently, we cannot search for exact matches, and the following approach was taken. First, each abstract's text was extracted from the meta-information files. Second, both the abstract's text and the article's main text were tokenized into sentences. Third, starting from the beginning of the main text, the first text

window with the same length as the abstract was compared to the abstract, and if the similarity was above 90%, the text window was labeled as the abstract and removed from the main text. If not, the window was shifted by one sentence, and the analysis repeated until the abstract was found or the window was shifted 20 sentences, whichever came first. Similarity was calculated using the FuzzyWuzzy Python package (version 0.17), which uses an edit distance algorithm, that is, counting the minimum number of operations (insert, delete, or substitute) required to transform one string to another (Jurafsky & Martin, 2008).

While abstracts were only available for 84% of articles, results indicate that compared to a base model including all the controls, adding style features for only the abstract does not increase predictive power (*adj. R*² = 0.218 vs. 0.217, *F* = 0.259, *p* = 0.985). This casts doubt on the notion that the effects are driven solely by abstracts

Specific function word types

Though our focus is in the effect of function words, overall, the relationships between specific function word types and citations are as follows: Personal Pronoun: *b* = −0.043, *p* < 0.001; Impersonal Pronoun: *b* = 0.088, *p* < 0.001; Article: *b* = −0.036, *p* < 0.001; Preposition: *b* = 0.016, *p* = 0.004; Auxiliary Verb: *b* = −0.023, *p* = 0.016; Adverb: *b* = 0.154, *p* < 0.001; Conjunction: *b* = 0.073, *p* < 0.001; Negation: *b* = −0.114, *p* < 0.001; Quantifier: *b* = 0.053, *p* < 0.001 ([Table S3](#)).

STUDY 2: EXPLORATORY ANALYSIS OF WRITING STYLES ASSOCIATED WITH IMPACT

While study 1 suggests that style shapes success, which particular ways of writing increase impact, and why?

Though the simplest place to start might seem like exploring the overall relationship between each type of function words and citations, this ignores a great deal of variation, both within function word type and due to context. Different types of personal pronouns (e.g., first-person vs. second-person), for example, have been shown to have different effects (e.g., on customer satisfaction, Packard et al., 2018). Consequently, assuming all function words in the same category have the same effect is likely incorrect. Further, given there are nine types of function words, comprehensively examining each in one paper would be challenging.

Consequently, we do something more exploratory. Rather than examining the overall relationships between a function word category and citations, which may be driven by various, even opposing effects, we explore specific contexts where our theorizing suggests that specific types of function words may have particular effects. We do not mean to suggest that these are the only effects of

function words, or the most important ones, but they are a place to start.

Specifically, we focus on three key areas. First, we explore auxiliary verbs. Auxiliary verbs like “had” or “has” can signal when something occurred (e.g., past like “had considered” or present like “has considered”), and we examine how such referencing different temporal perspectives relates to the prominence research achieves.

Second, we explore personal pronouns. We examine the role a particular type of personal pronouns (i.e., first-person pronouns like “I” or “we”) play in expressing personal voice.

Third, we examine prepositions and articles. These types of function words have been linked to cognitive complexity, and we examine how simplicity may be beneficial (or detrimental) depending on where it occurs.

Auxiliary verbs and temporal perspective

Auxiliary verbs (e.g., “has,” “had,” “is,” and “was”) modify content verbs and can signal their framing in time. When talking about the literature, for example, someone can talk about it in the past (e.g., it “had considered” the work on...) or present (e.g., it “has considered” the work on...).

While journal style guides, and academics themselves, commonly recommend describing research using the past tense (American Psychological Association, 2020; Bem, 2003; Nature, 2014), we suggest that present tense may lead to greater impact. Though a paper's content (i.e., theorizing, methods, and results) occurred in the past, using present tense may make that content seem more current and in the moment (Lieberman et al., 2007). This should make the ideas seem more relevant, applicable, and important to individual readers, which should make them more likely to cite the article. This, in turn, should increase an article's collective success (i.e., the number of citations it receives).

Method

To test this possibility, judges manually coded whether each auxiliary verb referenced the past (e.g., “had” or “was”) or present (e.g., “has” or “is”). Then, using the negative binomial model from the main analyses, and including all the control variables (i.e., non-language and content controls), we examine the relationship between past- and present-focused auxiliary verbs and the number of citations an article receives.

Results

Consistent with our theorizing, results suggest that past-focused language may reduce impact while present-focused language may increase it. While papers written with more past-focused auxiliary verbs are cited less

($b = -0.147, p < 0.001$), those written with more present-focused auxiliary verbs are cited more ($b = 0.031, p = 0.01$).

To test robustness, we also examine Linguistic Inquiry and Word Count's (Pennebaker et al., 2015) temporal orientation categories (i.e., focus past and focus present). These categories include 100 of words that discuss the past (e.g., added and appeared) or present (e.g., add and appear). Results are the same. More past-focused articles are cited less ($b = -0.125, p < 0.001$), while more present-focused articles are cited more ($b = 0.080, p < 0.001$; Table S6).

Personal pronouns and personal voice

Beyond temporal perspective, the use of personal pronouns may also shape citations. In particular, we examine personal pronouns that involve personal voice (i.e., first-person pronouns like “I” or “we”). While academic writing guides argue that authors should write in a manner that is distant, objective, and devoid of self-reference (e.g., avoiding words like “I” or “we”; Bem, 2003; Strunk Jr & White, 1999; c.f.; American Psychological Association, 2019), we suggest that using such first-person pronouns may sometimes be beneficial.

Pronouns suggest who is driving the action. If someone says “we found X,” for example, it suggests that *they* found something. Alternatively, using passive voice (e.g., “X was found”) or a different subject (e.g., “this research found”) distances the authors from the action. Consistent with this idea, employees who used more first-person pronouns (“I” or “we”) while helping customers seemed more involved or agentic (Packard et al., 2018).

Building on this notion, we suggest that taking personal ownership of arguments, hypotheses, findings, and contributions (e.g., “we reveal” vs. “the present research reveals”) that are empirically supported may make the authors seem more knowledgeable, increasing the perceived authority of the research. This should lead individual readers to be more likely to cite the piece, and, in turn, lead articles that use such language to be more successful overall. Taking personal ownership of methods and results (e.g., “we asked participants to do X” vs. “participants did X”), however, may make methodological choices seem more subjective. Consequently, whether personal voice is beneficial or not may depend on where it is used (i.e., front end vs. methods and results).

Method

To test this possibility, we use manual coding and machine learning to separate each article into three segments, (1) front end (i.e., literature review and theorizing), (2) methods and results, and (3) general discussion or conclusion (see Appendix S1 for more detail). Then, we use the negative binomial model

from the main analyses including all the control variables (i.e., non-language and content controls) to test the relationship between first-person pronouns and citations across different article segments (controlling for all other function word categories in each of those segments). This sheds light on the relationship between first-person pronouns in a given article portion (e.g., the front end) and citations, controlling for the presence of that feature in other portions (i.e., the conclusion).

Results

As expected, and in contrast to some prior suggestions, using personal voice does not always seem detrimental. While papers written with more self-referencing function words (i.e., first-person pronouns) in the empirical section are cited less ($b = -0.029$, $p = 0.016$), papers whose front ends are written with more first-person pronouns are cited *more* ($b = 0.076$, $p < 0.001$). First-person pronouns had a different relationship with citations across different sections (middle section $\times$ first-person pronouns: $b = -0.062$, $p < 0.001$; conclusion section $\times$ first-person pronouns: $b = -0.058$, $p < 0.001$). Results are similar for usage of “I” and “we.” Thus, while some writing manuals are more comfortable with usage of “we” than “I” it does not seem like there is a distinction between them. The relationship was also positive in the concluding section, albeit not significantly so ($b = 0.007$, $p = 0.29$). A multiple function word measure used to identify more personable writing style (i.e., analytic thinking; Jordan et al., 2019) shows similar results.

Further, if our theorizing about the value of removing personal ownership from the methods and results section is correct, impersonal pronouns may also be important. Impersonal pronouns (e.g., “it” or “that”) remove a personal actor and thus may be valuable for distancing the authors from the methodological choice and findings. Consistent with this notion, articles which used more impersonal pronouns in the middle empirical section (i.e., methods and results) were cited more ($b = 0.080$, $p < 0.001$; Table S7).

Articles, prepositions, and simplicity

Simplicity may also be linked to the success of ideas. While academic ideas are often quite complex (Gray, 2021; Metoyer-Duran, 1993), communicating them more simply should make them easier to understand, remember, and act or build on (Chater, 1999; Fogg, 2009). This should increase the likelihood that readers cite an article, which, in turn, should increase the article's collective impact (i.e., number of citations it achieves). We suggest this may be particularly important when ideas

are first being explained (e.g., in the beginning of an article where authors are laying out their thinking).

To test these possibilities, we explore two types of function words linked to cognitive complexity: articles and preposition (Biber et al., 2011; Pennebaker et al., 2003; Pennebaker & King, 1999). Grammatical articles add complexity by asking readers to make distinctions between a single case or a class of something (e.g., *the* car vs. *a* car) while prepositions do so by introducing novel or conditional linkages among nouns, pronouns, or phrases (e.g., growth *despite* inflation; ate more candies *except* when).

Method

To test these possibilities, similar to the analyses involving personal voice, we use the negative binomial model from the main analyses. We include all the control variables (i.e., non-language and content controls) and test the relationship between articles and prepositions and citations across different article segments (controlling for all other function word categories in each of those segments).

Results

Consistent with our theorizing, results indicate that simplicity is linked to citations. Papers that use fewer articles ($b = -0.034$, $p < 0.001$) or prepositions ($b = -0.015$, $p < 0.001$) in the front end are cited more. Traditional readability measures show similar effects. Though we did not have an explicit prediction about the concluding section, effects there are similar, albeit weaker ($b_{\text{Articles}} = -0.007$, $p < 0.001$; $b_{\text{Prepositions}} = -0.004$, $p = 0.17$). There is some evidence that the relationship between articles and citations varied by sections (conclusion section $\times$ articles: $b = 0.008$, $p = 0.06$; middle section $\times$ articles: $b = 0.003$, $p = 0.55$). Similarly, there is some evidence that the relationship between prepositions and citations varied by sections (middle section $\times$ prepositions: $b = 0.011$, $p = 0.06$; conclusion section $\times$ prepositions: $b = 0.0004$, $p = 0.93$).

That does not mean simplicity is always good, though. In the methods and results, for example, complexity may sometimes be useful or even required. Indeed, the cost of complexity seems to weaken ($b_{\text{Articles}} = -0.002$, $p = 0.66$) or even reverses ($b_{\text{Prepositions}} = 0.027$, $p < 0.001$) in the middle section where methods or results are discussed (Table S8).

Discussion

Overall, an exploratory analysis sheds some initial light on categories of function words linked to greater impact. Specifically, consistent with our theorizing, results

suggest that auxiliary verbs (through their role in communicating time), personal pronouns (through their role in personal voice), and articles and prepositions (through shaping complexity) are all associated with how many citations research achieves.

Importantly, though, as the results illustrate, these relationships can depend on the context. While papers with simpler front ends are cited more, for example, simplicity is less useful in the methods and results.

There are certainly other reason function words may shape communication's impact, but hopefully these initial results spark future research.

STUDY 3 AND 4: EXPERIMENTAL EVIDENCE

While natural language processing of tens of thousands of articles suggests that linguistic style may shape impact, one could still wonder whether the effects are truly causal. After all, maybe there is some unobserved factor that is driving both linguistic style and citations.

Controlling for factors like author fame, institution, article topics, other linguistic features, and over a dozen other measures casts doubt on this possibility, but to provide an even stronger test, we conduct two simple experiments. In each, we manipulate one of the function word categories linked to increased citations in Study 2 and explore its causal impact. Specifically, we manipulate personal voice through the presence of pronouns (Study 3) and temporal perspective through verb tense (Study 4).

STUDY 3: PERSONAL VOICE

Study 3 tests our suggestions that using personal pronouns to take ownership of a finding may make research seem more important and impactful. Further, it tests whether this occurs, at least in part, because personal voice makes authors seem more insightful or knowledgeable.

Method

Participants ($N = 185$, recruited through Amazon Mechanical Turk) were randomly assigned to condition (i.e., Control or Personal Voice) in a between-subjects design. They were told they would read some information about academic research.

All participants read a brief one-sentence research description, and the only difference between conditions was whether or not the description included personal voice. In the control [personal voice] condition, participants read that “Results [We] reveal an even more cost-efficient way to produce clean energy.”

All participants then completed the dependent measures, the importance of the research (i.e., “how important is this research”) and its likely impact (i.e., “how impactful do you think this research will be”), both on 7-point scales (1 = not at all, 7 = extremely). To test the hypothesized process, participants also completed a two-item measure of how knowledgeable the author seemed about the topic (i.e., “how knowledgeable do you think the author is about this research topic?”, “how insightful is the author when it comes to this issue?”; $r = 0.86$, $p < 0.001$).

Results

As predicted, and consistent with the results of the field data, using personal voice led people to think the study was more important ($M_{\text{Personal Voice}} = 6.15$ vs. $M_{\text{Control}} = 5.76$, $F(1,183) = 5.35$, $p = 0.022$). Using personal voice also led them to think the study would be more impactful ($M_{\text{Personal Voice}} = 5.97$ vs. $M_{\text{Control}} = 5.39$, $F(1,183) = 8.16$, $p = 0.005$).

Finally, as predicted, mediation analysis (PROCESS model 4; Hayes, 2018) confirmed that perceived author knowledge mediated the effect of personal voice on the study's importance (indirect effect = 0.15, 95% CI [0.05, 0.26]) and impact (indirect effect = 0.19, 95% CI [0.06, 0.34]). Using personal voice made the author seem more knowledgeable about the topic ($b = 0.27$, $SE = 0.09$, $t = 3.01$, $p = 0.003$), which made the study seem more important ($b = 0.56$, $SE = 0.06$, $t = 10.21$, $p < 0.001$) and impactful ($b = 0.69$, $SE = 0.07$, $t = 10.65$, $p < 0.001$).

STUDY 4: TEMPORAL PERSPECTIVE

Study 4 tests our suggestion that talking about a finding using present rather than past tense auxiliary verbs makes research seem more important and impactful. Further, it tests whether this is driven, at least in part, by making the research seem more current.

Method

Participants ($N = 234$, recruited through Prolific) were randomly assigned to condition (i.e., Present Tense or Past Tense auxiliary verbs) in a between-subjects design.

All participants read a brief one-sentence research description, and the only difference between conditions was the auxiliary verb tense used. In the past [present] tense condition, participants read that “This drug had [has] achieved a 15% reduction in cancer cells.”

All participants then completed the dependent measures from Study 3. To test the hypothesized process,

participants also completed a measure of how current the findings are (i.e., “how current is this finding?” 1 = Not at all current; 7 = Extremely current).

Results

As predicted, and consistent with the field data, using present tense led people to think the study was more important ($M_{\text{Present}} = 5.97$ vs. $M_{\text{Past}} = 5.60$, $F(1, 232) = 5.36$, $p = 0.021$) and would be more impactful ($M_{\text{Present}} = 5.69$ vs. $M_{\text{Past}} = 5.20$, $F(1, 200) = 7.79$, $p = 0.006$).

Further, as predicted, mediation analysis (PROCESS model 4; Hayes, 2018) found that perceived currentness of the findings mediated the effect of personal voice on the study's importance (indirect effect = 0.24, 95% CI [0.09, 0.41]) and impact (indirect effect = 0.22, 95% CI [0.08, 0.38]). Using personal voice made the research seem more current ($b = 0.55$, $SE = 0.16$, $t = 3.38$, $p < 0.001$) which boosted the study's perceived importance ($b = 0.43$, $SE = 0.06$, $t = 7.41$, $p < 0.001$) and impact ($b = 0.40$, $SE = 0.06$, $t = 6.20$, $p < 0.001$).

Discussion

Experimental evidence provides further evidence for the impact of writing style. Consistent with the field data, small shifts in function words (e.g., replacing the word “results” with the pronoun “we”) increased the perceived importance and expected impact of academic research. An additional experiment (Study 5, see Appendix S1) demonstrates that the results of Study 4 extend beyond auxiliary verbs to verb tense more generally. Describing research results using present rather than past tense (i.e., a drug lowers rather than lowered cancer risk) led participants to think the study was more important and the findings would be more impactful.

By directly manipulating these language features, we cast doubt on alternative explanations based on unobservables or correlated content and underscore function words' causal impact. By keeping the content identical, and simply manipulating style, we further demonstrate style's importance.

GENERAL DISCUSSION

Language shapes the effectiveness of communication. It shapes whether ads persuade, salespeople sell, and customer service agents are seen as helpful. It shapes whether word of mouth influences behavior, leaders are effective, or academic research has impact.

But while content certainly shapes the success of ideas, the present work finds that style also plays an important role. Even in academic research, where writing is often seen as merely a disinterested way to communicate

truth, writing style helps explain the impact (i.e., number of citations) ideas achieve. Style words (i.e., “function words”) explain 4–11% of the overall variance explained and 11–27% of language's impact on citations. If communication style shapes impact even in a content-focused domain like academic research, its effect in other domains might be even larger.

The results also suggest specific writing styles that may increase impact. While journal style guides often suggest using past tense, results suggest that using the present tense (e.g., we *theorize* instead of *theorized*) may be more beneficial. While some have suggested that first-person voice is bad because it distracts from the paper's content (Bem, 2003), we find that there are times for taking personal credit for writing (i.e., front end) and others for letting the paper's content stand on its own (i.e., methods and results). And while academic ideas are often complex, explaining ideas simply may be important, particularly in a paper's front end.

Contributions and implications

These findings have a number of contributions and implications. First, they speak to how to communicate academic research. Communicators often think that using complex language will make them seem more intelligent (Oppenheimer, 2006). Further, peer-reviewed research often adopts a dry, dense, and impersonal style that can be sometimes challenging to read and understand (Freeling et al., 2019; Ruben, 2016). But while authors across many disciplines have intermittently theorized about what counts as “better writing” (Bem, 2003; McCloskey, 1985; Mensh & Kording, 2017), little work has actually tested these suggestions.

Our results suggest that a few relatively simple stylistic shifts may help boost academic research's impact. Using first-person pronouns (e.g., “we” suggest) to own arguments, directions, or contributions may, for example, be helpful. When discussing the evidence for one's ideas, however, avoiding self-reference may make choices seem less subjective. Using a simpler style (i.e., fewer articles and prepositions), particularly at the outset, may also be beneficial. And while style guides commonly recommend past tense, this may be misguided. Using present tense rather than past (e.g., “we find” rather than “we found”) may increase impact by making the ideas seem more current and relevant.

Along these lines, the results suggest *when* particular language features might play a greater role. Content is often divided into multiple parts or portions. The goals of a presentation's introduction, for example, might be quite different than the goals of the middle (e.g., to garner interest versus inform). Similarly, the introduction of an academic paper may have different goals than the methods and results. Consequently, simple suggestions like “write simply” may not only miss nuance but be

detrimental in some cases. Indeed, our results suggest that writing too simply in the methods and results may decrease impact (see Markowitz, 2019 for the relationship between different types of linguistic complexity and grant funding). By dividing papers into sections, we explore not only whether style matters, but when.

This work also contributes to the literature on cultural selection (Akpınar & Berger, 2015; Heath et al., 2001; Kashima, 2008; Norenzayan et al., 2006; Schaller & Crandall, 2003). While cross-cultural research has demonstrated how culture affects individual-level psychological processes (Markus & Kitayama, 1991), the reverse is also true. Psychological processes shape the norms, practices, and content that make up culture. In this instance, linguistic style may make articles easier to read and comprehend, make methodological choices seem more objective, or make findings seem more important, novel, or generalizable. These aspects, in turn, should shape each reader's likelihood of citing the article, and thus, the prominence research achieves.

Future research

As with any investigation, these findings raise a number of interesting questions for future research. Work might examine the relative impact of style versus content in different domains. One could argue that academic writing is relatively constrained. Papers have norms or style guides, and given the topical content, there may be constraints on language features like pronoun use or certain types of grammatical articles. Consequently, the effects might be even larger in areas like advertising or public speaking where language is less constrained.

The inferences generated by the style variations considered might also vary across different age groups, cultures, or education levels. And while we controlled for some aspects of content (i.e., the topics discussed and other linguistic features), one could argue that the quality or novelty of ideas being presented has not been taken into account. Though it is possible that controlling for the topical mixture gets at this to some degree, it may not get at it fully. While we are not aware of any measures of idea quality, hopefully future advances will allow this to be studied in greater detail. Idea novelty might be able to be captured by how similar content is to related work, but doing so remains non-trivial.

Present tense may also increase persuasion (see Packard et al., 2023). Saying a book “is” rather than “was” interesting made others more interested in reading it, for example, and online reviews noting that a beach “does” rather than “did” have great atmosphere made readers think they would like it more. Using present tense may make communicators seem more confident, which may make others more likely to follow their suggestions.

Future research might also delve into other ways language shapes an idea's impact. Expressing more certainty,

for example, may be beneficial (because it increases the perception that a phenomenon is true) or detrimental (if it seems unwarranted; Willis et al., 2011). Ancillary analyses in our data, for example, suggest that LIWC's clout measure (Kacewicz et al., 2014), which captures the degree to which people speak with confidence and a sense of certainty, is associated with greater citations, though it seems that this relationship is driven mostly by the middle section of the paper.

The degree to which articles are written using more of a narrative style may also play a role. Ancillary analyses in our data, for example, suggest that lower scores on LIWC's analytical thinking measure (Pennebaker et al., 2014), which indicate more narrative thinking, is associated with greater citations. Further, rather than being restricted solely to the language in the introduction, this relationship also exists, albeit weakly, through the middle and concluding section of papers.

Using more familiar language may help if it makes content easier to read. Compared to using definite articles (i.e., “the,” which specifies a singular, identified member of the type, e.g., violating *the norm of* ...) using indefinite articles (i.e., “a” or “an” which means that any member of that type is being discussed, e.g., violating *a norm of* ...) may make content seem broader and generally applicable, which may increase citations. Language's impact may also vary by discipline. While jargon may generally decrease readability (Pennebaker et al., 2015; Warren et al., 2021), and thus citations, it may increase impact in disciplines where it is seen as a signal of credibility.

Finally, this work highlights the value of using natural language processing to understand human behavior (see Berger & Packard, 2022; Boyd & Schwartz, 2021 for recent reviews). Most activities people do on a daily basis involve language. People produce language through writing emails, making presentations, talking to friends, and posting content online. They consume language through watching movies, listening to songs, and reading books, news articles, and other types of content. Even human thought involves language.

As these examples highlight, language serves two key functions. It both (1) *reflects* aspects of the people, organizations, and socio-cultural contexts that produce it and (2) *impacts* the people and audiences that consume it. Consequently, analyzing language can shed light on both language producers and language recipients. It can deepen understanding around differences between people, organizations, and cultures. And it can provide insight into persuasion, effective communication, and why ideas catch on. By quantifying features of language, automated text analysis will hopefully unlock a range of interesting insights.

ACKNOWLEDGEMENTS

Thanks to the Wharton Behavioral Lab for help running the experiments.

CONFLICT OF INTEREST STATEMENT

Authors declare no conflict of interest

ORCID

Reihane Boghrati  <https://orcid.org/0000-0002-5281-1310>

Grant Packard  <https://orcid.org/0000-0002-8765-1994>

REFERENCES

- Akpinar, E., & Berger, J. (2015). Drivers of cultural success: The case of sensory metaphors. *Journal of Personality and Social Psychology*, 109(1), 20–34.
- American Psychological Association. (2019). *First-person pronouns*. <http://apastyle.apa.org/style-grammar-guidelines/grammar/first-person-pronouns>
- American Psychological Association. (2020). *Publication manual of the American psychological association*. American Psychological Association.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania.
- Bem, D. J. (2003). Writing the empirical journal. In J. M. Darley, M. P. Zanna, & H. L. Roediger, III (Eds.), *The complete academic: A practical guide for the beginning social scientist* (2nd ed.). American Psychological Association.
- Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science*, 22(7), 891–893.
- Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2), 192–205.
- Berger, J., & Packard, G. (2022). Using natural language processing to understand people and culture. *American Psychologist*, 77(4), 525.
- Berry-Blunt, A. K., Holtzman, N. S., Donnellan, M. B., & Mehl, M. R. (2021). The story of “I” tracking: Psychological implications of self-referential language use. *Social and Personality Psychology Compass*, 15(12), e12647.
- 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.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993–1022.
- Boyd, R. L., & Schwartz, H. A. (2021). Natural language analysis and the psychology of verbal behavior: The past, present, and future states of the field. *Journal of Language and Social Psychology*, 40(1), 21–41.
- Chater, N. (1999). The search for simplicity: A fundamental cognitive principle? *The Quarterly Journal of Experimental Psychology: Section A*, 52(2), 273–302.
- Chen, C. (2012). Predictive effects of structural variation on citation counts. *Journal of the American Society for Information Science and Technology*, 63(3), 431–449.
- Chung, C., & Pennebaker, J. W. (2007). The psychological functions of function words. *Social Communication*, 1, 343–359.
- Cristianini, N., Shawe-Taylor, J., et al. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.
- Fogg, B. J. (2009). A behavior model for persuasive design. In *Proceedings of the 4th international conference on persuasive technology* (pp. 1–7). ACM New York.
- Fox, C. W., Burns, C. S., Muncy, A. D., & Meyer, J. A. (2016). Gender differences in patterns of authorship do not affect peer review outcomes at an ecology journal. *Functional Ecology*, 30(1), 126–139.
- Freeling, B., Doubleday, Z. A., & Connell, S. D. (2019). Opinion: How can we boost the impact of publications? Try better writing. *Proceedings of the National Academy of Sciences*, 116(2), 341–343.
- Gray, D. E. (2021). *Doing research in the real world*. Sage.
- Hamermesh, D. S., Johnson, G. E., & Weisbrod, B. A. (1982). Scholarship, citations and salaries: Economic rewards in economics. *Southern Economic Journal*, 49, 472–481.
- Hamermesh, D. S., & Schmidt, P. (2003). The determinants of econometric society fellows elections. *Econometrica*, 71, 399–407.
- Harzing, A. W. (2007). Publish or perish. <https://harzing.com/resources/publish-or-perish>
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (Vol. 2). Springer.
- Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, 85(1), 4–40.
- Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology*, 81(6), 1028–1041.
- 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.
- Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, 22(1), 39–44.
- Jordan, K. N., Sterling, J., Pennebaker, J. W., & Boyd, R. L. (2019). Examining long-term trends in politics and culture through language of political leaders and cultural institutions. *Proceedings of the National Academy of Sciences*, 116(9), 3476–3481.
- Jurafsky, D., & Martin, J. H. (2008). *Speech and language processing: An introduction to speech recognition, computational linguistics and natural language processing*. Prentice Hall.
- Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33(2), 125–143.
- Kaggle. (2017). US baby names, Version 2. <https://www.kaggle.com/kaggle/us-baby-names>
- Kashima, Y. (2008). A social psychology of cultural dynamics: Examining how cultures are formed, maintained, and transformed. *Social and Personality Psychology Compass*, 2(1), 107–120.
- Lakoff, G., & Johnson, M. (1999). *Philosophical in the flesh: Cognitive science brings to philosophy the embodied mind, the cognitive unconscious, and metaphorical thought*. University of Chicago Press.
- Lawson, M. A., & Matz, S. C. (2022). Saying more than we know: How language provides a window into the human psyche. In *The psychology of technology: Social science research in the age of big data* (pp. 45–85). American Psychological Association. <https://doi.org/10.1037/0000290-003>
- Lieberman, N., Trope, Y., McCrea, S. M., & Sherman, S. J. (2007). The effect of level of construal on the temporal distance of activity enactment. *Journal of Experimental Social Psychology*, 43(1), 143–149.
- Markowitz, D. M. (2019). What words are worth: National Science Foundation grant abstracts indicate award funding. *Journal of Language and Social Psychology*, 38(3), 264–282.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253.
- McCloskey, D. (1985). Economical writing. *Economic Inquiry*, 23(2), 187–222.
- Mensh, B., & Kording, K. (2017). Ten simple rules for structuring papers. *PLOS Computational Biology*, 13, e1005830.

- Metoyer-Duran, C. (1993). The readability of published, accepted, and rejected papers appearing in College & Research Libraries. *College & Research Libraries*, 54(6), 517–526.
- Nature. (2014). *Writing scientific papers*. Nature: Education <https://www.nature.com/scitable/ebooks/english-communication-for-scientists-14053993/126083980/>
- Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive Science*, 30(3), 531–553.
- Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology*, 20(2), 139–156.
- Packard, G., Berger, J., & Boghrati, R. (2023). How verb tense shapes persuasion. *Journal of Consumer Research*, ucad006. <https://doi.org/10.1093/jcr/ucad006>
- Packard, G., Moore, S. G., & McFerran, B. (2018). (I'm) happy to help (you): The impact of personal pronoun use in customer-firm interactions. *Journal of Marketing Research*, 55(4), 541–555.
- Pennebaker, J. W. (2011). Using computer analyses to identify language style and aggressive intent: The secret life of function words. *Dynamics of Asymmetric Conflict*, 4(2), 92–102.
- Pennebaker, J. W. (2017). Mind mapping: Using everyday language to explore social & psychological processes. *Procedia Computer Science*, 118, 100–107.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. The University of Texas at Austin.
- Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PLoS One*, 9(12), e115844.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6), 1296–1312.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54(1), 547–577.
- Pinker, S. (2014). Why academics stink at writing. *The Chronicle of Higher Education*, 61(5). <https://www.chronicle.com/article/why-academics-stink-at-writing/>
- Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.
- Ruben, A. (2016). How to read a scientific paper. *Science*, 20. <https://www.sciencemag.org/careers/2016/01/how-read-scientific-paper>
- Schaller, M., & Crandall, C. S. (2003). *The psychological foundations of culture*. Psychology Press.
- Smart, S., & Waldfogel, J. (1996). *A citation-based test for discrimination at economics and finance journals*. National Bureau of Economic Research.
- Stremersch, S., & Verhoef, P. C. (2005). Globalization of authorship in the marketing discipline: Does it help or hinder the field? *Marketing Science*, 24(4), 585–594.
- Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: Drivers of article impact. *Journal of Marketing*, 71(3), 171–193.
- Strunk, W., Jr., & White, E. B. (1999). *The elements of style*. Longman.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.
- Topaz, C. M., & Sen, S. (2016). Gender representation on journal editorial boards in the mathematical sciences. *PLoS One*, 11(8), e0161357.
- Van Wesel, M., Wyatt, S., & ten Haaf, J. (2014). What a difference a colon makes: How superficial factors influence subsequent citation. *Scientometrics*, 98(3), 1601–1615.
- Warren, N. L., Farmer, M., Gu, T., & Warren, C. (2021). Marketing ideas: How to write research articles that readers understand and cite. *Journal of Marketing*, 85, 42–57.
- Willis, D. L., Bahler, C. D., Neuberger, M. M., & Dahm, P. (2011). Predictors of citations in the urological literature. *BJU International*, 107(12), 1876–1880.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Boghrati, R., Berger, J., & Packard, G. (2023). Style, content, and the success of ideas. *Journal of Consumer Psychology*, 00, 1–13. <https://doi.org/10.1002/jcpy.1346>