

PSYCHOLINGUISTICS

The narrative arc: Revealing core narrative structures through text analysis

Ryan L. Boyd^{1*}, Kate G. Blackburn², James W. Pennebaker²

Scholars across disciplines have long debated the existence of a common structure that underlies narratives. Using computer-based language analysis methods, several structural and psychological categories of language were measured across ~40,000 traditional narratives (e.g., novels and movie scripts) and ~20,000 nontraditional narratives (science reporting in newspaper articles, TED talks, and Supreme Court opinions). Across traditional narratives, a consistent underlying story structure emerged that revealed three primary processes: staging, plot progression, and cognitive tension. No evidence emerged to indicate that adherence to normative story structures was related to the popularity of the story. Last, analysis of fact-driven texts revealed structures that differed from story-based narratives.

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

INTRODUCTION

Narratives permeate all facets of human life. We tell stories to teach cultural norms, to entertain, and to help create shared perspectives. We construct stories to make meaning out of past events and to create new worlds and possibilities for ourselves and others. The creation of narratives is a fundamental human behavior that spans recorded history.

Narrative theorists, philosophers, literary scholars, linguists, and countless other disciplines have long attempted to identify the quintessential properties of narratives (1–3). Among them, there is a general consensus that narratives typically follow a predefined path that includes a beginning, middle, and end. Such theoretical frameworks primarily invoke the notion that the broad-brush characteristics of a story provide a narrative's structure while granting the possibility of infinite variations on the specific content around this structure. Put another way, most narrative frameworks assume that the structure and content of stories are independent—for example, two mystery novels need not have the same characters or settings to have similar plot dynamics.

The task of finding objective and quantifiable markers of narrative structure is daunting. Early empirical studies relied on small numbers of texts using human coders. Research has provided mixed evidence for common plot structures, recurring cultural symbols, and character archetypes (4–6). In recent years, however, automated text analysis methods, together with increasing access to growing numbers of texts, have given rise to promising new ways to quantify the underlying and universal properties of stories (7–11). For example, approaches that tap the coherence and cohesion of text, such as Coh-Metrix (12), can mathematically compute the similarity of word meanings from sentence to sentence and across different parts of a text. Cohesion scores, then, can infer the degree to which a text remains on the same general topic as it unfolds.

Other recent, content-related approaches attempt to track how linguistic markers of emotion unfold over the course of stories. A recent study reported an analysis of over 1700 digitized novels and other texts, where common emotional peaks and valleys were systematically quantified and clustered (13). In total, the researchers identified six common affective trajectories (i.e., “emotional arcs”)

that emerged from stories, each corresponding to different plot archetypes. For example, a consistent rise in positive emotion words over the course of a story was indicative of a “rags-to-riches” storyline, whereas a decline in positive emotion words was indicative of tragedy.

The research on emotional arcs raises broader questions about narrative structure. Narratives are much more than affective experiences—they are presented in such a way as to support information about the characters, conflict and resolution, and various other story-specific features. The 19th century scholar Gustav Freytag famously depicted several narrative elements as points on a generalized narrative arc [commonly referred to as the “dramatic arc” (14)]. Freytag argued that exposition, or the building of the story's scene, laid the groundwork for a narrative's structure. As a story moves forward, action between characters increases and, ultimately, peaks at the top of the narrative arc: the story's climax. Subsequently, a decline in conflict prompts characters to transition toward the denouement or the resolution. Freytag described these narrative elements as something akin to specific events within a story—relatively discrete destinations that appear at fixed points in a narrative.

At its essence, Freytag's framework suggests three primary processes in the unfolding of a story. The first is the narrator setting the stage and establishing the context for the story. Once the elements of the story are established, plot progression begins through the movement of characters across time and space with increasing interactions among them. Moreover, the focal point of a story is the central conflict or cognitive tension that the characters must grapple with and ultimately resolve. Freytag, of course, is not the only person to identify these three elements. Aristotle, for example, hypothesized that these narrative elements were strung together by storytellers to maintain the audience's engagement and memory of stories (15). A frustration of modern literary and social scholars, however, has been the difficulty of objectively and reliably quantifying the narrative process across a wide array of stories.

Is it possible to use text analytic methods to capture markers of staging, plot progression, and cognitive tension? Critically, these three narrative processes should be present to some degree or another in all stories, regardless of the actual content within a narrative. In analyzing stories, then, the building blocks of narrative should be manifest in relatively content-free words. In English, content-heavy phrases, sentences, and paragraphs are supported largely by syntax-related words called “function” words. Whereas content words include nouns, regular verbs, and most adjectives and adverbs, function

¹Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK. ²Department of Psychology, The University of Texas at Austin, Austin, TX 78712-1043, USA.

*Corresponding author. Email: r.boyd@lancaster.ac.uk

words are typically short connector words such as pronouns, prepositions, articles, conjunctions, negations, auxiliary verbs, and non-referential adverbs (e.g., “so” and “really”).

Although there are fewer than 200 commonly used function words in English, they account for 50 to 60% of all words that we say, write, hear, or read (16). Multiple studies conducted over the last two decades suggest that function words require social skills and awareness to use and to interpret. Imagine finding a note on the street:

“I am with them now and she is there. Back soon.”

The words are common and the note is syntactically coherent, but the ultimate meaning behind the message is elusive. Using only function words, the note refers to people, places, and time that only the author and the intended reader share knowledge of. We can infer the author’s relationship with the intended reader (they are probably close) and possibly other social psychological attributes of the author at rates higher than chance. Multiple studies have documented that function words can be used to identify social psychological factors including people’s age, sex, relative status in an interaction, honesty, formality, intelligence, and emotional state (17).

In the study of narrative, then, function words should serve as valuable markers of a story’s progression. When a story begins, the author and the reader are strangers. To set up the story, the reader needs a great deal of contextual information: Who are the characters? What are the relationships among them? When and where is the story taking place? At the beginning, then, the author must signal concrete labels, names, and other identifying clues for the characters, places, and objects in the story; importantly, the author must also connect these dots by elaborating on their interrelations. In providing the necessary background, the author must necessarily use high rates of prepositions and articles (the mansion was next to the lake, below a bluff, by the road)—words that are inherently information-structural (18). Once the reader becomes familiar with the context, the author can later refer to the mansion as “it” or “her home” or perhaps not at all. Once the plot gets moving, there should be a large increase in pronouns, auxiliary verbs, and other function words and a corresponding drop in articles and prepositions.

Psychometrically, prepositions and articles behave differently from other function words. In an analysis of college admissions essays of over 25,000 students, articles and prepositions were positively correlated +0.25, while the other function word categories were inter-correlated +0.21. More notable was that the overall average correlation between articles and preposition with the other function words averaged –0.30 (19). Inspections of essays that were high in articles/prepositions found that they were analytical, formal, and hierarchically structured in terms of how information was presented. Those essays that scored high in pronouns, auxiliary verbs, and so on were found to be far more character and action oriented, as well as more informal. However, this distinction was not typological—all essays contained some elements of both, and what was found to differ was the continuum along which any given author used information-structural language versus story-driving language. In a sense, stories have both an analytic and a narrative component. At the beginning, the author must set the stage with a logical and structured landscape. Once the general landscape has been established, the plot can move forward through the action of the characters and interceding events (20).

The third dimension of Freytag’s model concerns the cognitive tension or climax of a story. A narrative traditionally revolves around a vexing, unresolved problem. Once the stage has been set and the narrative begins unfolding, the main characters face central obstacles

or frustrations to their goals or challenges to their worldviews. It is during these challenges that cognitive tension forms in a narrative—characters must actively process their scenarios, resolve conflict, and form new ways of understanding their worlds. In the same way that staging and plot progression are both present to some degree throughout a narrative, the cognitive tension that arises from conflict is theorized to be present throughout a story, even to some degree at or near the very beginning (21). Over the years, several researchers have discovered that a group of cognitive processing words (e.g., “think,” “believe,” “understand,” and “cause”) reflects the sense-making process that people engage in while working through a conflict or challenge in their life (22, 23). Further, when people are uncertain of an answer to a complex problem, cognitive process words are used at higher rates than if the person is absolutely certain of the answer (24).

Building on the extensive literature demonstrating the relationship between function words, cognitive words, and psychological processes involved in language, we set out to better understand the intersection between these widely understood verbal-behavioral psychological markers and their congruence with traditional theories of narrative structure. The primary purpose of the current research was to determine if it is possible to quantify the three narrative processes by relying on the analysis of function and cognitive process words across different types of narratives. Three central research questions emerged.

RQ1: Do different types of traditional stories share similar narrative arcs reminiscent of Freytag’s three processes?

By aligning language features with narrative theory, we have built an exploratory model to evaluate whether a common, normative structure does exist across narratives. On the basis of this model, narrative processes develop in predictable ways. According to our narrative arc framework, we predict the following.

1) Staging: Narrators establish the basic background or set the stage by establishing the names, locations, and relationships among characters and places at the beginning of a story. To do so, they must invoke prepositions and articles at high rates, which are used to construct comprehensible situations to the reader (25). After the introduction of essential background information, authors will need fewer explicit concept references and, over the course of the story, the use of prepositions and articles should diminish in their relative frequency.

2) Plot regression: Once background information is established, the audience has a shared understanding of context with the narrator, allowing the storyteller to move the plot forward. A story’s momentum should be detectable through pronouns, auxiliary verbs, and other common function words. Words such as pronouns and connectives are essential not only for narrative cohesion and coherence as a story progresses (26, 27) but also to move the plot through important experiential processes that are cued to the audience (28).

3) Cognitive tension: As stories reach a climax, the characters and narrators work through the issues they face in an increasingly active fashion, cognitively speaking. The relative amount of cognitive tension, or “working-through,” will be reflected in cognitive process words (29, 30). Consistent with Freytag’s model, the cognitive tension phase of the story should typically peak during the middle-to-late parts of stories.

As described in greater detail below, the current project initially focuses on three primary corpora of traditional narratives: novels, short stories, and even shorter stories that were written spontaneously

by a large number of internet users. If narratives do have standard, normative structures, we expect that each of the three very different corpora would exhibit a parallel unfolding of structural patterns.

RQ2: Do popular/highly rated stories exhibit different narrative structures from less popular stories?

The essence of a narrative is the story it tells. We have all had the experience of loving a movie or novel that some of our friends detest. Intuitively, the differences of opinion tend to surround the underlying themes rather than on the staging or plot progression. If true, one could make an argument that the narrative structure could be similar between popular and unpopular movies or books. Alternatively, less popular movies and books may be less liked due to their unconventional (or overly conventional) structure. The second research question sought to analyze the transcribed dialogue of movies and romance novels, both of which have been rated by a large number of audience members.

RQ3: Do different types of narratives—such as those in the popular coverage of science or even TED talks—have the same structure as traditional narratives?

Definitions of narratives and stories are highly fluid. The most traditional assumption is that a narrative contains a series of connected events with a beginning, middle, and end. Generally, a narrative is considered a superordinate construct composed of a series of stories or events that could be imaginary or factual. Whereas novels and short-form fiction are narratives, the factual accounts of scientific discoveries are likely less story-like but are still narratives nonetheless. The third research question seeks to establish some of the boundary conditions of the arc of narrative to determine the degree to which scientific stories may adhere to the same narrative structure as more traditional forms of storytelling.

MATERIALS AND METHODS

At the outset, it was not clear how strong the arc of narrative signal might be. Given that no previous scholars have successfully identified linguistic patterns that indicate a normative narrative structure, we sought to analyze a reasonably large sample of stories using a relatively conservative word count strategy. A more detailed account of the methods is available in the Supplementary Materials. See also the accompanying website (<https://arcofnarrative.com>) for demonstrations of the arc of narrative approach to selected novels and user-provided stories; with this tool, users are able to create visualizations of the data, measure narrative processes in their own texts, and view additional information about the methods and this research.

Corpora

For RQ1, three initial corpora were collected that spanned genres, authorship attributes, and centuries. The traditional stories included 2523 fiction novels from the Gutenberg corpus, 2092 short stories from various online sources, and 14,419 brief stories written by internet users in response to a standardized thematic apperception test [TAT; (31)] from a website maintained by the third author.

Additional corpora were collected for subsequent research questions, including a large corpus of transcriptions of movie dialogue ($N = 19,970$) provided by OpenSubtitles.org, supplemented with IMDb ratings for each film. Self-published romance novels ($N = 639$) col-

lected from Smashwords.com were analyzed, including community ratings of each text. Last, two science-related corpora included 18,312 *New York Times* science articles published between 1989 and 2017 and 2226 transcriptions of TED talks. Last, a nonscience corpus of 1580 Supreme Court majority, concurring, and dissenting opinions were included. Complete details about text inclusion criteria and data cleaning procedures are presented in the Supplementary Materials.

Text analytic approach

The analytic strategy involved quantifying the rates of function and cognitive words across each text. Function word rates associated with staging (articles and prepositions) were calculated separately from those associated with plot progression (pronouns, auxiliary verbs, negations, conjunctions, and nonreferential adverbs). The sum of each of the two function word categories for each text was divided by the total number of words in the text yielding the percentage of total words. Cognitive tension words were based on the cognitive process dictionary from the text analysis program Linguistic Inquiry and Word Count [LIWC2015; (32)]. The development and psychometrics of the narrative process dictionary is described in the Supplementary Materials, along with the specific variations that we tested to ensure robustness of the dictionary. The total number of words, word stems, and lemmatized words varied considerably across the three dictionaries: staging, 75 words; plot progression, 448 words; and cognitive tension, 393 words.

To capture the unfolding of the three theorized processes over the course of the narratives, each text was split into five equal-sized segments. The five-segment approach was initially chosen to roughly correspond with the five traditional “stages” of a narrative (i.e., exposition, rising action, climax, falling action, and denouement). As described in the Supplementary Materials (section C), we obtained conceptually consistent results regardless of whether the texts were partitioned into different numbers of segments, ranging from 3 to 10. The language dimensions within each segment were analyzed separately, yielding 15 narrative measurements per story (3 narrative processes \times 5 text segments), which allowed for the modeling of the dynamics of narrative processes as they progressed through each story.

Base rates of the three dictionaries varied considerably by genre; all scores were standardized within each text separately for each narrative dimension to focus on the relative, story-internal shape of each narrative dimension across texts. As an example, recall that staging is the sum of all articles and prepositions, expressed in percentage of total words. If, in a particular text, the rates of staging (that is, percentage of articles and prepositions) across the five segments changed from 7.4 to 3.3%, the within-text standardized scores for that text would be as follows:

	Segment				
	1	2	3	4	5
Raw scores	7.4	8.2	5.5	6.0	3.3
Standardized	0.70	1.10	-0.30	-0.04	-1.40

Very broadly, the data were analyzed in two ways: traditional hypothesis testing methods and predictive modeling. For most research questions, we used simple mixed design analyses of variance (ANOVAs) to test the effects of the three narrative processes over the five segments across the types of narratives under investigation. Using orthogonal polynomial contrasts over the five segments, we tested the

degree to which the language variables adhered to linear and quadratic functions. Given such large datasets, we only discuss main effects and interactions that are statistically significant at $P \leq .001$. Instead of focusing on probability levels, we rely primarily on estimates of percent variance accounted for using the partial eta squared (η_p^2) statistic.

While ANOVA tests are primarily reported here due to their ease of interpretation for a wide readership, we conducted additional, more sophisticated predictive modeling analyses to more accurately capture and subsequently understand the nature of each narrative process's shape. These additional analyses centered around polynomial functions and cross-validation methods; tests of model fitting are outlined in the Supplementary Materials (sections G and H). Extensive, additional information on the quantification and standardization of narrative processes within and between texts is presented in the Supplementary Materials.

RESULTS AND DISCUSSION

The three overlapping research questions address the parameters of the arc of narrative. The first question attempts to establish the degree to which more traditional fictional stories have normative arcs concerning staging, plot progression, and cognitive tension. The second research question extends the idea into two narrative fiction corpora—movie scripts and romance novels—which have been rated by large numbers of audience members. The goal is to determine whether highly rated stories have different narrative arcs than those that are less liked. The final question seeks to determine whether stories other than traditional narratives adhere to the same underlying structures.

RQ1: Do common, normative narrative process arcs exist for traditional stories?

Recall that the three corpora of traditional narratives included well-known novels published by established writers, short stories written by authors of varying experience, and brief TAT stories that online visitors to a psychology website wrote in response to a standardized ambiguous drawing.

An initial 3 (corpus) \times 3 (narrative process: staging, plot progression, cognitive tension) \times 5 (story segment) mixed design ANOVA yielded a modest overall segment main effect, accounting for 3% of the variance, $\eta_p^2 = 0.031$, and a narrative process by segment inter-

action accounting for over 8% of the variance, $\eta_p^2 = 0.085$. All other effects, including interactions with corpus, accounted for less than 1% of the variance.

As can be seen in the three panels of Fig. 1, the graphs of each of the narrative processes are similar across the three corpora. Because of the overall process by segment interactions, separate corpus by segment analyses were computed for each process.

Further, the variance associated with the segment effects were partitioned into linear and U-shaped quadratic effects. The left panel depicts the usage of words associated with staging across the five segments of the texts. Consistent with predictions, the use of staging-related words is highest at the beginning of the story and drops shortly thereafter as the story unfolds. A simple corpus by linear segment ANOVA for the staging dimension yielded a modest main effect for segment, $\eta_p^2 = 0.028$.

The findings for the plot progression narrative process also resulted in patterns predicted by the arc of narrative model. As can be seen in the middle panel, words associated with plot progression were lowest in the beginning parts of the texts and sharply increased thereafter. A simple corpus by segment ANOVA for the plot progression dimension yielded modest main effects for the orthogonal linear and U-shaped quadratic segment effects (linear term $\eta_p^2 = 0.069$; quadratic term $\eta_p^2 = 0.017$).

Last, the unfolding of the cognitive tension process can be seen in the right panel of Fig. 1. As predicted, the cognitive process words were used most midway in the narrative. The only noteworthy effect was the quadratic segment main effect, $\eta_p^2 = 0.018$. As with the other narrative processes, the differences between the corpora were inconsequential.

As noted earlier, we conducted additional, more detailed analyses using linear and polynomial ordinary least squares regression and involved training models via a competitive 10-by-10-fold cross-validation experiment to better capture the “true” shape of each narrative process for each corpus. Briefly described, these analyses consistently showed that all three narrative processes show clear, strong, normative quadratic arcs that are consistent in terms of both shape and magnitude across all corpora. Put another way, for each analysis described above, we urge the reader to keep in mind that each shape is definitively nonlinear in nature—consistent with intuitions that can be drawn from the figures below, no narrative dimension was

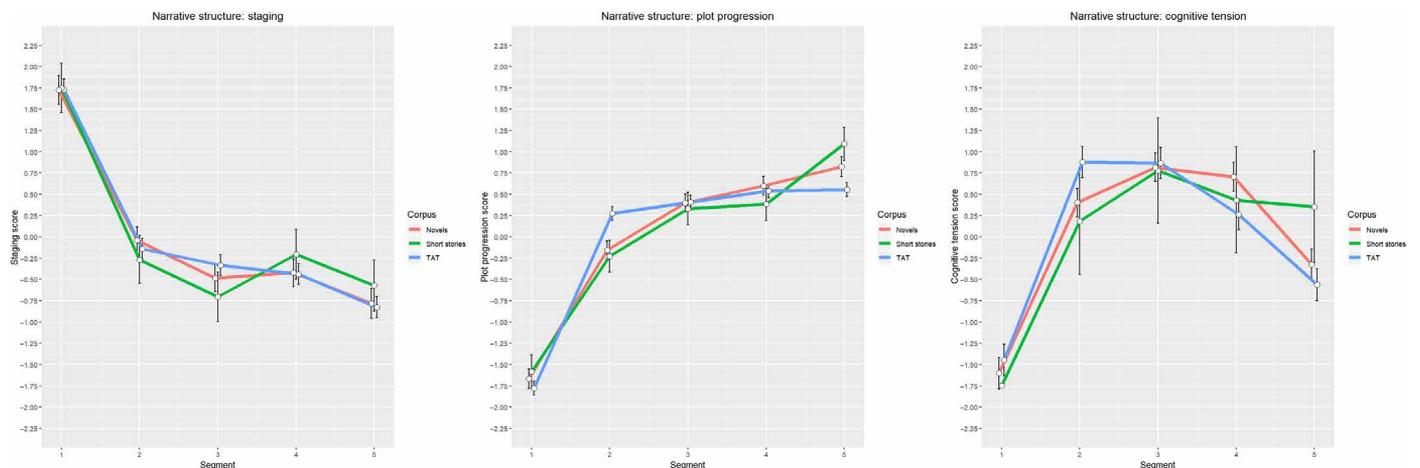


Fig. 1. Normative means and standard errors of each narrative process, by corpus. Means and error bars rescaled to illustrate the typical narrative from each genre.

best represented as a linear function. Comprehensive details on these analyses are presented in the Supplementary Materials.

The importance of the first research question was to demonstrate that the theorized processes existed, were consistent with the Freytag hypotheses, and were reasonably consistent across three types of stories that varied considerably in length, writing experience of the authors, the intended audiences, and the amount of time spent thinking about and writing the stories. Although the overall effect sizes were modest, analyses showed that the shapes and magnitudes of each narrative process were highly consistent across all corpora.

RQ2: Are there differences in narrative structure between popular and less popular stories?

The second research question sought to determine whether the quality or, perhaps, popularity of stories depended, in part, on their narrative structure. For these analyses, we relied on two additional text corpora—romance novels and the dialogue of movies—that could be used to address RQ2 in addition to serving as replication samples for results found for RQ1. Romance novels and movies were selected due to the fact that they constitute traditional narratives like those used for RQ1 with the added benefit of availability of subjective ratings made by readers of the books and viewers of the movies. All texts were prepared/analyzed in a manner identical to that described for RQ1.

As a literary form, romance novels date back to ancient Greece and, 2000 years later, account for one of today's top-selling fiction genres (33). Unlike the broad range of stories analyzed in RQ1, most romance novels are relatively similar in length, content, and structure (34). The romance novel corpus included 639 self-published books retrieved from Smashwords.com, a free e-book and online publishing platform with a catalog of over 400,000 books. Only recent books written in English that had reader ratings were included. The ratings ranged from 1 to 5 stars and were averaged across raters for each book. Although the ratings were highly skewed toward five stars, we arbitrarily classified the books into three categories: low ratings (with 3.9 or fewer stars, $N = 149$ books), mid-level ratings (for 4 to 4.9 stars, $N = 348$ books), or perfectly rated (5 stars, $N = 142$ books).

Initially, an overall 3 (ratings group: high, mid-level, low) \times 3 (narrative process) \times 5 (segment) mixed-effects ANOVA was computed on the narrative process scores. The analyses revealed a large segment main effect ($\eta_p^2 = 0.123$) and narrative process \times segment interaction ($\eta_p^2 = 0.445$). The only remaining effect was a modest ratings group by narrative process \times segment interaction ($\eta_p^2 = 0.019$). To dissect the interactions, separate ratings by orthogonal linear and quadratic segment analyses were computed for each narrative process. More details on the analyses are included in the Supplementary Materials.

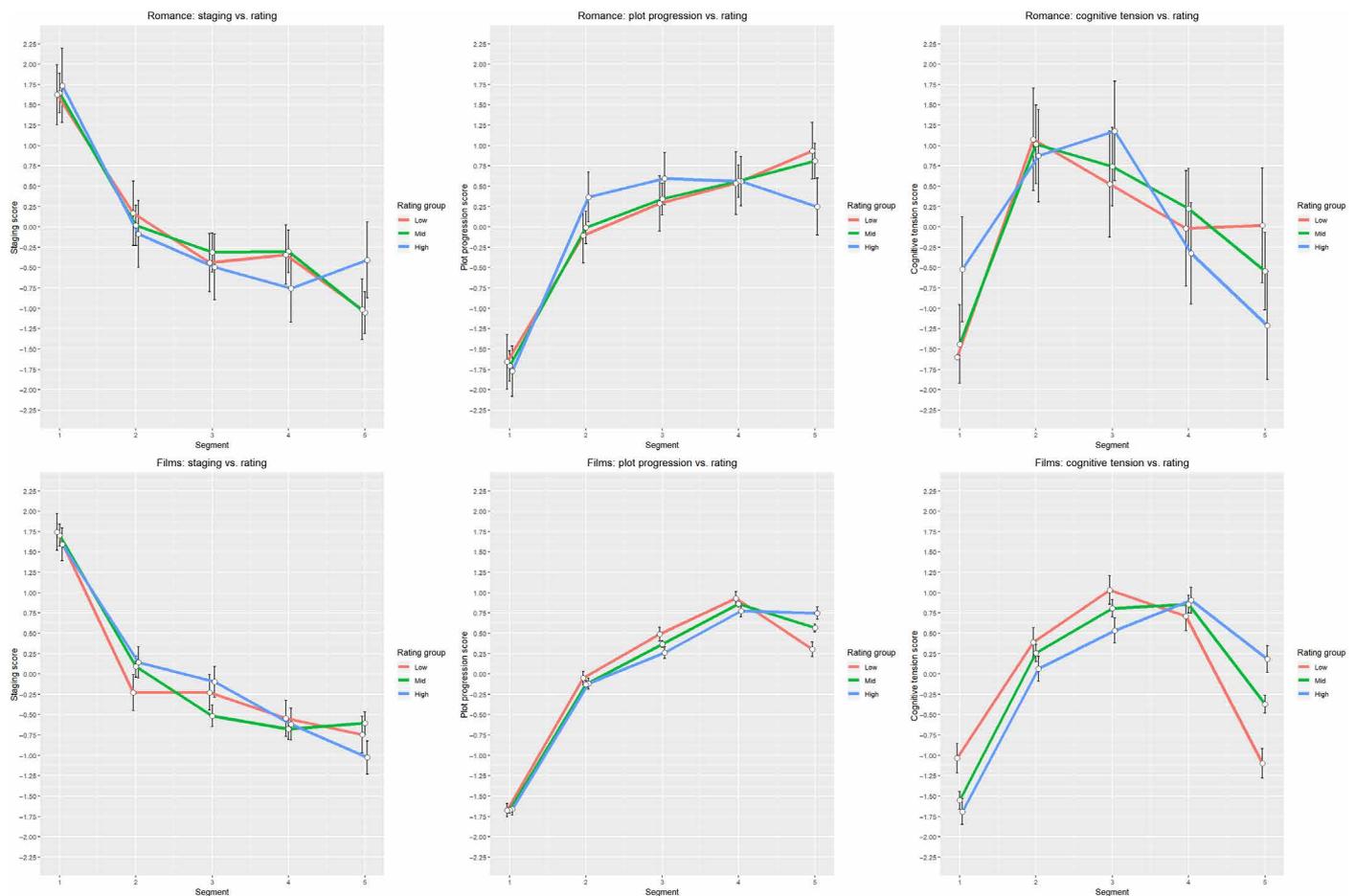


Fig. 2. Means of each narrative dimension for romance novels (top row) and film dialogue (bottom row), plotted separately by rating group (low, medium, and high). Means and error bars rescaled to illustrate the typical narrative from each genre.

The three narrative processes for each of the ratings groups for the romantic novels are displayed along the top panel of Fig. 2. The results for staging ($\eta_p^2 = 0.241$; left panel), plot progression ($\eta_p^2 = 0.298$; middle panel), and cognitive tension ($\eta_p^2 = 0.091$; right panel) reveal similar patterns to the traditional story corpora from RQ1. The linear term effect sizes for staging and plot progression were far higher than the RQ1 findings; the cognitive tension quadratic effect was comparable. The most important finding, however, was the failure to find any meaningful book rating by segment or narrative process interactions.

Overall, the analyses of the romance novels indicate that they have the same general narrative processes as traditional novels, short stories, and TAT stories. However, there is no hint that popular stories have different structures than those rated more negatively by readers. The study suffers from two possible problems. The first concerns the authors and ratings: Most of the writers were not professionals. Further, the raters are not known and many may be acquaintances of the authors. Second, the sample size of the romance novels was considerably smaller than the RQ1 corpora.

A conceptually similar replication of RQ2 was undertaken with the dialogue of the 18,074 movies that had a minimum of 100 user ratings from IMDb at the time of data collection. Each English language movie was rated by registered IMDb users along a 10-point scale ranging from 1 being “terrible” to 10 being “excellent.” IMDb does not release its actual rating algorithm, which is weighted by metrics associated with users. However, IMDb reports that the rating algorithms are consistent across all movies and are intended to reduce the likelihood of a small group of people exploiting the rating system (35).

Movie ratings were positively skewed ($M = 6.33$, $SD = 1.20$) but with fewer than 5% rated 8.0 or higher. To capture narrative arcs across a wide range of preferences, three groups of movies were created: low ratings (≤ 5.0 , $N = 2546$), mid-range (5.1 to 7.0, $N = 10,204$), and high (greater than 7.0, $N = 5324$). As with the romance novels, an initial 3 (rating group) \times 3 (narrative process) \times 5 (segment) between-within repeated-measures ANOVA was run followed by separate rating group by segment ANOVAs for each of the narrative processes to disentangle any interactions.

As can be seen along the bottom panel of Fig. 2, the same general patterns emerged for the three narrative processes for the movie corpus as with all other corpora. In addition, the differences in narrative processes as a function of movie ratings were minimal. The overall multivariate ANOVA yielded robust segment effects, $\eta_p^2 = 0.157$, and a narrative process by segment interaction, $\eta_p^2 = 0.208$. Separate ANOVAs by narrative process yielded a small segment effect for the staging process, $\eta_p^2 = 0.039$, and a stronger effect for plot progression, $\eta_p^2 = 0.213$. The segment cognitive tension effects were small for the linear term effect, $\eta_p^2 = 0.012$, and modest for the quadratic effect, $\eta_p^2 = 0.057$. No meaningful main effect or interaction emerged with rating group.

The results of all RQ2 analyses found no evidence that people’s ratings of narrative quality or popularity were related to narrative structure. Across both corpora, there was no hint of an interaction with ratings for either the narrative process or story segment. Put another way, the narrative structure for all three narrative process was unrelated to how much the audience liked each story.

RQ3: Do other standard story formats have similar narrative structures?

To summarize thus far, the five corpora that have been analyzed were all fictional stories that share the same basic narrative processes: (i)

high rates of words that reflect stage setting at the beginning, subsequently dropping over the course of the story; (ii) relatively low use of plot progression words at the beginning that increase steadily until leveling off close to the end of the story; and (iii) an inverted-U use of words related to cognitive tension peaking midway in the story. Our empirical findings directly confirm what many narrative theorists have speculated about for thousands of years. The mean of these three processes is depicted in Fig. 3. As can be seen in this figure, we find that the three narrative dimensions do appear to be consistent with past theorizing on narrative structure: Normatively speaking, staging, plot progression, and cognitive tension appear unfold in a manner much like Freytag’s description of narratives (14).

The distinction between narratives and stories is often blurred. Whereas the types of stories examined in RQ1 and RQ2 can be thought of as falling under a conventional definition of fictional narratives, other types of stories not only exist but also are quite commonplace. For example, news reporting is commonly referred to as a “news story,” despite the format and goals of news stories being markedly different from a fiction novel or a fairy tale (36). Our primary question for RQ3, then, is to determine whether the three narrative processes behave in the same way for narratives that do not fall under the umbrella of traditional fictional stories.

To answer the third research question, we focus on three conceptually different narrative types: 28,664 *New York Times* articles reporting on developments in the fields of science and technology, 2226 TED talks distributed by the TED website, and 1580 Supreme Court majority, concurring, and dissenting opinions by all 12 justices that sat on the bench from 2002 to 2009. Newspaper articles are examples of informative writing intended to relay facts to readers to help them understand the causes and possible effects of a particular event or experience. TED talks are highly rehearsed narratives generally

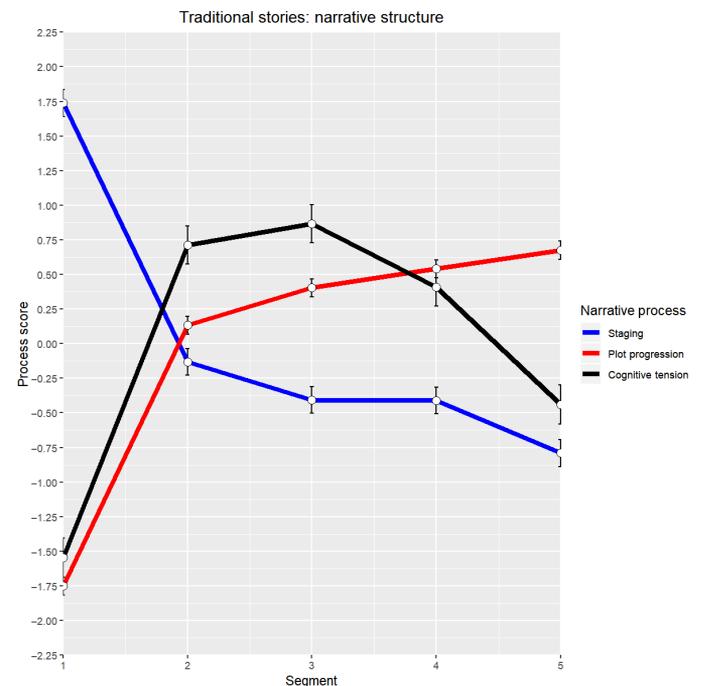


Fig. 3. Typical narrative processes across the five traditional story corpora used for RQ1 and RQ2 (novels, short stories, TATs, romance novels, and movie dialogue).

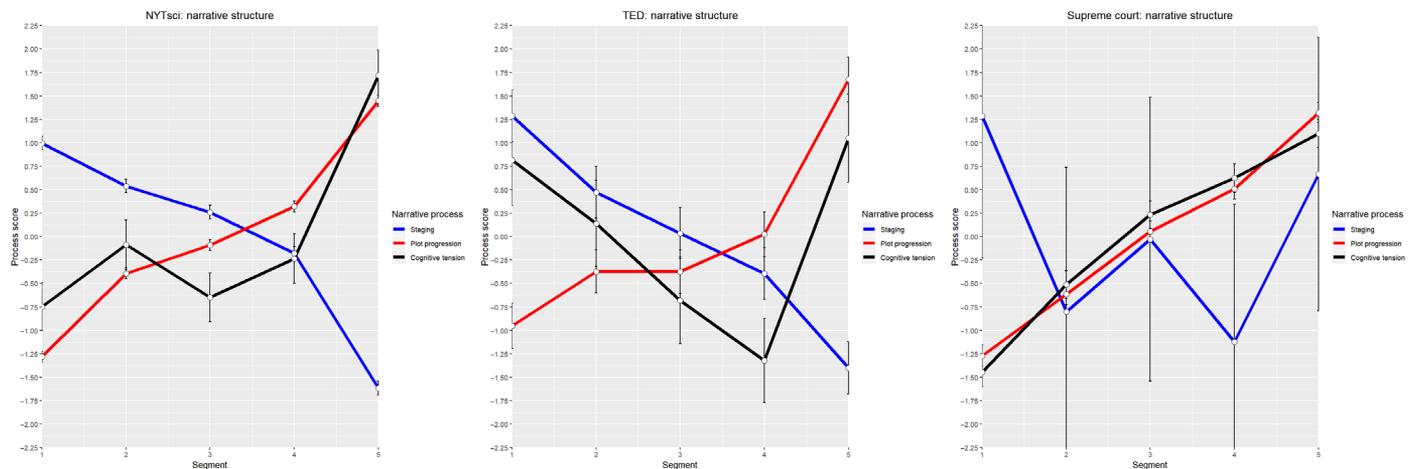


Fig. 4. Unfolding narrative structure by genre. Note: Graphs represent the three narrative processes across the five segments for *New York Times* articles (left), TED talks (center), and Supreme Court opinions (right). Means and error bars rescaled to illustrate the typical narrative from each genre.

based on the speakers' research or experiences that have clear implications for technology, education, or design (37). Supreme Court decisions are a form of expository writing whereby judges analyze a particular case, lay out its connection to legal precedence, and provide a logical opinion and its implications.

Separate analyses were conducted on each corpus. We encourage the reader to pause and compare the narrative processes from traditional stories (Fig. 3) with the process found for the *New York Times*, TED talks, and Supreme Court decisions as shown in Fig. 4.

The three corpora reveal relatively similar patterns to the original story corpora for both staging and plot progression. For the *New York Times* and TED talks, the staging lines drop in a pattern comparable to those in RQ1 and RQ2 ($\eta_p^2 = 0.098$ and 0.063). The Supreme Court arguments, on the other hand, exhibited no staging changes. For plot progression, the *New York Times*, TED talks, and Supreme Court corpora all exhibit consistent increases over time ($\eta_p^2 = 0.150$, 0.068 , and 0.372) in ways virtually identical to the traditional stories.

The cognitive tension dimension revealed the most notable contrast between the nonfiction texts and the traditional stories. Recall that cognitive processing words are used most when people are trying to understand or work through a difficult problem. In the fictional stories in RQ1 and RQ2, the most unanswered and puzzling questions were typically in the middle of the narrative. In the *New York Times* and TED talks, but not the Supreme Court, linear term effects were small or nonexistent ($\eta_p^2 = 0.005$, 0.001 , and 0.226). Unlike the stories in RQ1 and RQ2, the quadratic effects were negligible ($\eta_p^2 = 0.002$, 0.016 , and 0.008). The quadratic effect for the TED talks was small, yet suggested a U-shaped function rather than an inverted-U.

The findings from the RQ3 analyses suggest that different types of narratives may have their own unique narrative arc. Whereas traditional stories are built around one or more essential questions or puzzles that are eventually solved, more evidence-based narratives serve very different functions. For example, the field of journalism has historically trained writers to construct newspaper articles using a relatively standard format: The essential facts of an article are provided in the first one to two paragraphs, followed by supporting information, and concluding with relevant questions or possible implications for the future. TED talks, not unlike scientific publications, begin with background information, followed by methods and findings,

ending with questions and implications for the future. Neither newspaper articles nor TED talks typically have explicit plots around which the narratives revolve. In both cases, the unresolved questions are typically at the end of the narrative.

Supreme Court opinions have a somewhat different language structure with a decidedly more formal writing style (38, 39). In most cases, court opinions address facts and conclusions that have previously been decided in lower court decisions. The audience of most decisions is typically composed of lower court judges and other legal scholars. The crux of most court decisions is linking the facts of the case with previous legal decisions and existing legal doctrine with which the audience is presumed to already know.

SUMMARY AND CONCLUSIONS

We sought to determine whether the narrative processes and structures that have been theorized by historical scholars such as Aristotle and Freytag could be captured using psychologically informed measures of language. Across ~40,000 traditional narratives, we found strong, highly consistent evidence of these structures. Stage setting, plot progression, and cognitive tension exhibited coherent patterns of unfolding across genres, authorship attributes, and story lengths. Moreover, each narrative dimension unfolded in theoretically consistent ways: staging tends to occur at its highest at the beginning of a story, followed by a rise in plot progression, paired with a rise and fall in cognitive tension around the middle-to-late parts of a story. To our knowledge, this is the first set of studies that have provided empirical support for classical theories of the underlying dimensions, and structure, of traditional narratives. In addition, our findings rest at the intersection of psychological and narrative theory, bridging the gap between the social psychological, cognitive, and literary scholarship.

Three research questions guided the current project. The first sought to determine if traditional stories evidenced normative patterns of function and cognitive word use that tapped into fundamental properties of an unfolding story: setting the stage, plot progression, and cognitive tension. Across five large corpora of traditional stories, linguistic markers of each demonstrated highly consistent, normative patterns of progression. The second research question found no clear differences between highly rated and poorly rated movies and

romance novels in terms of their narrative structure. The third research question found that fictional stories had their own distinctive narrative arc compared with the narrative structures of newspaper articles, TED talks, and Supreme Court opinions.

Several fundamental questions are raised by the current project. Although the discovered effects are consistent across a large number of texts, are highly statistically significant, and account for a respectable amount of variance, the underlying meaning of each of the narrative processes will require greater scrutiny. For example, the process of setting the stage has been defined by the use of articles and prepositions; however, articles and prepositions do not set the stage on their own. Rather, they are low-level, structural markers of the authors' need to label objects, people, and events and connect them to each other, time, and space. Similarly, the function word categories that signal plot progression are a mix of intercorrelated language dimensions that suggest a shared understanding of named people, places, and events (through the use of pronouns) and signal ongoing activity (through auxiliary verbs).

At the heart of the issue is answering why consistent patterns of narrative processes emerge across such diverse stories. It is possible that the structure found within the current research provides an optimal system for delivering narrative information. That is, an audience may be able to navigate the illustration of a story more readily/usefully under a given set of narrative process structures. For example, if the layout of a narrative does not include information about the setting of the story early on, readers may find it difficult to understand a character's motivations, goals, and behaviors as the plot moves forward.

From an evolutionary perspective, the structure of storytelling may provide a crucial way for people (or different groups) to share information. Numerous language acquisition and developmental studies have demonstrated that while children quickly assign names for objects and people, ascribing action to these objects proves more difficult (40, 41). Similarly, early stages of language in primates adhere to a similar structure by showing that a speaker's language is object oriented, and only later is action given to those objects (42). The optimal structure of storytelling, then, may originate from a natural inclination to first define objects/people and then assign action. In other words, the ways in which a story's information is processed may follow particular parameters to which narratives have evolved to adhere, in general. This need not be a unidirectional process—the way that we understand stories perhaps influences the process by which we create stories, imposing a particular order and structure on language during storytelling.

The effect sizes for the narrative process models vary considerably from corpus to corpus. Those narratives that are the most highly constrained in terms of their topic and structure (e.g., romance novels, novels, and Supreme Court opinions) accounted for much more overall variance than the less uniform styles and topics of the other corpora. Although the effect sizes are modest across datasets, the narrative patterns are unlikely to be visible to the human eye without the assistance of computers, and the structures may be seen as an underlying germ or genomic blueprint upon which narratives are developed and permuted. While a generalized narrative structure was found to exist across tens of thousands of stories, great variability between narratives did exist, suggesting that a strong creative element influences a narrative's deviation from the norm.

We are only now beginning to understand the structure and function of stories from an empirical, scientific perspective. Several questions remain for future research. While a generalized narrative

structure has been identified, we do not yet know how deviation from normative narrative structure might reflect things like creativity, abstractness, or even genre variability. Similarly, future work should more deeply explore the psychological function of such structures, much in the way that recent research has explored psychologically central topics of self-narratives (10). In addition, the degree to which these methods and structures may apply to other forms of language, such as social interactions or instructional discourse, remains unknown. As our understanding of narratives progresses, these questions and others may be addressed with increasingly refined and complex methodologies.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/32/eaba2196/DC1>

REFERENCES AND NOTES

1. Aristotle, L. J. Potts, *Aristotle on the Art of Fiction: An English Translation of Aristotle's Poetics* (Cambridge Univ. Press, 1968).
2. J. Franklin, *Writing for Story: Craft Secrets of Dramatic Nonfiction by a Two-Time Pulitzer Prize Winner* (Plume, 1994).
3. N. Frye, *Anatomy of Criticism: Four Essays* (Princeton Univ. Press, 2000).
4. J. Campbell, *The Hero with a Thousand Faces* (Pantheon Books, 1949).
5. W. Labov, J. Waletzky, *Narrative Analysis: Oral Versions of Personal Experience* (John Benjamins, 1997), vol. 7.
6. V. Propp, *Morphology of the Folktale* (University of Texas Press, 1968).
7. A. C. Graesser, M. Singer, T. Trabasso, *Constructing Inferences During Narrative Text Comprehension* (American Psychological Association, 1994), vol. 101.
8. H. Li, A. C. Graesser, M. Conley, Z. Cai, P. I. Pavlik, J. W. Pennebaker, A new measure of text formality: An analysis of discourse of Mao Zedong. *Discourse Process*. **53**, 205–232 (2016).
9. J.-B. Michel, Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, J. P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M. A. Nowak, E. L. Aiden, Quantitative analysis of culture using millions of digitized books. *Science* **331**, 176–182 (2011).
10. K. C. McLean, M. Syed, M. Pasupathi, J. M. Adler, W. L. Dunlop, D. Drusturp, R. Fivush, M. E. Graci, J. P. Lilgendahl, J. Lodi-Smith, D. P. McAdams, T. P. McCoy, The empirical structure of narrative identity: The initial Big Three. *J. Pers. Soc. Psychol.* (2019).
11. P. S. Dodds, E. M. Clark, S. Desu, M. R. Frank, A. J. Reagan, J. R. Williams, L. Mitchell, K. D. Harris, I. M. Kloumann, J. P. Bagrow, K. Megerdooomian, M. T. McMahon, B. F. Tivnan, C. M. Danforth, Human language reveals a universal positivity bias. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2389–2394 (2015).
12. A. C. Graesser, D. S. McNamara, M. M. Louwse, Z. Cai, Coh-Matrix: Analysis of text on cohesion and language. *Behav. Res. Methods Instrum. Comput.* **36**, 193–202 (2004).
13. A. J. Reagan, L. Mitchell, D. Kiley, C. M. Danforth, P. S. Dodds, The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Sci.* **5**, 31 (2016).
14. G. Freytag, E. J. MacEwan, *Technique of the Drama: An Exposition of Dramatic Composition and Art; An Authorized Translation from the 6th German Edition* (Scholarly Press, 1960).
15. M. Mack, *The Norton Anthology of World Masterpieces* (Norton, ed. 5, 1985), vol. 1.
16. J. W. Pennebaker, *The Secret Life of Pronouns: What Our Words Say About Us* (Bloomsbury, 2011).
17. C. Chung, J. Pennebaker, in *Social Communication*, K. Fiedler, Ed. (Frontiers of Social Psychology, Psychology Press, 2007), pp. 343–359.
18. K. N. Jordan, J. Sterling, J. W. Pennebaker, R. L. Boyd, Examining long-term trends in politics and culture through language of political leaders and cultural institutions. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3476–3481 (2019).
19. J. W. Pennebaker, C. K. Chung, J. Frazee, G. M. Lavergne, D. I. Beaver, When small words foretell academic success: The case of college admissions essays. *PLOS ONE* **9**, e115844 (2014).
20. N. R. Norrick, *Conversational Narrative* (John Benjamins Publishing Company, 2000).
21. M.-L. Ryan, The modal structure of narrative universes. *Poetics Today* **6**, 717–755 (1985).
22. B. Kleim, A. B. Horn, R. Kraehenmann, M. R. Mehl, A. Ehlers, Early linguistic markers of trauma-specific processing predict post-trauma adjustment. *Front. Psychol.* **9**, 645 (2018).
23. J. M. Caspari, T. L. Raque-Bogdan, C. McRae, T. L. Simoneau, S. Ash-Lee, K. Hultgren, Posttraumatic growth after cancer: The role of perceived threat and cognitive processing. *J. Psychosoc. Oncol.* **35**, 561–577 (2017).
24. M. A. Khawaja, F. Chen, C. Owen, G. Hickey, in *Human-Computer Interaction—INTERACT 2009* (Lecture Notes in Computer Science, Springer, 2009), pp. 485–489.

25. D. G. Morrow, Spatial models, prepositions, and verb-aspect markers. *Discourse Process*. **13**, 441–469 (1990).
26. M. W. Lovett, M. Dennis, J. E. Newman, Making reference: The cohesive use of pronouns in the narrative discourse of hemidecorticate adolescents. *Brain Lang*. **29**, 224–251 (1986).
27. S. Kleijn, H. L. W. P. Maat, T. J. M. Sanders, Comprehension effects of connectives across texts, readers, and coherence relations. *Discourse Process*. **56**, 447–464 (2019).
28. F. Hartung, M. Burke, P. Hagoort, R. M. Willems, Taking perspective: Personal pronouns affect experiential aspects of literary reading. *PLOS ONE* **11**, e0154732 (2016).
29. J. G. Bohanek, R. Fivush, Personal narratives, well-being, and gender in adolescence. *Cogn. Dev.* **25**, 368–379 (2010).
30. P. J. Bauer, E. N. Stark, A. F. Lukowski, J. Rademacher, D. L. V. Abbema, J. K. Ackil, Working together to make sense of the past: Mothers' and children's use of internal states language in conversations about traumatic and nontraumatic events. *J. Cogn. Dev.* **6**, 463–488 (2005).
31. H. A. Murray, *Thematic apperception test* (Harvard University Press, Cambridge, MA, US, 1943).
32. J. W. Pennebaker, R. L. Boyd, K. Jordan, K. Blackburn, *The Development and Psychometric Properties of LIWC2015* (The University of Texas at Austin, 2015).
33. U.S. adult fiction book sales 2017 | Statistic. *Statista*; <https://www.statista.com/statistics/730316/adult-fiction-unit-sales/>.
34. P. Regis, *A Natural History of the Romance Novel* (University of Pennsylvania Press, 2007).
35. IMDb | Help; <https://help.imdb.com/article/imdb/track-movies-tv/ratings-faq/G67Y87TFYYP6TWAV#accurate>.
36. B. Kovach, *The Elements of Journalism: What Newspapers Should Know and the Public Should Expect* (Three Rivers Press, 2014).
37. Our organization; <https://www.ted.com/about/our-organization>.
38. R. A. Posner, Judges' writing styles (and do they matter). *U. Chi. L. Rev.* **62**, 1421 (1995).
39. R. J. Owens, J. P. Wedeking, Justices and legal clarity: Analyzing the complexity of U.S. Supreme Court opinions. *Law Soc. Rev.* **45**, 1027–1061 (2011).
40. D. Gentner, *Why Nouns are Learned Before Verbs: Linguistic Relativity Versus Natural Partitioning* (257, Illinois University, Center for the Study of Reading, 1982), pp. 1–86.
41. D. Gentner, Language as cognitive tool kit: How language supports relational thought. *Am. Psychol.* **71**, 650–657 (2016).
42. M. Tomasello, *Constructing a Language: A Usage-Based Theory of Language Acquisition* (Harvard Univ. Press, 2003).
43. D. C. McClelland, J. W. Atkinson, R. A. Clark, E. L. Lowell, *The Achievement Motive* (Irvington, 1953).
44. D. C. McClelland, *Power: The Inner Experience* (Irvington, 1975).
45. F. D. Schönbrodt, B. Hagemeyer, V. Brandstätter, T. Czirkmanti, P. Gröpel, M. Hennecke, L. S. F. Israel, K. Janson, N. Kemper, M. Köllner, P. M. Kopp, A. Mojzisch, R. Müller-Hotop, J. Prüfer, M. Quirin, B. Scheidemann, L. Schiestel, S. Schulz-Hardt, L. Sust, C. Zygarr, O. C. Schultheiss, Measuring implicit motives with the Picture Story Exercise (PSE): Databases of expert-coded German stories, pictures, and updated picture norms. *PsychArchives*, (2020).
46. J. W. Pennebaker, R. J. Booth, R. L. Boyd, M. E. Francis, *Linguistic Inquiry and Word Count: LIWC2015* (Pennebaker Conglomerates, 2015).
47. E. Kacewicz, J. W. Pennebaker, M. Davis, M. Jeon, A. C. Graesser, Pronoun use reflects standings in social hierarchies. *J. Lang. Soc. Psychol.* **33**, 125–143 (2013).
48. R. L. Boyd, in *Data Analytics in Digital Humanities*, S. Hai-Jew, Ed. (Springer International Publishing, 2017), pp. 161–189.
49. R. L. Boyd, J. W. Pennebaker, Did Shakespeare write Double Falsehood? Identifying individuals by creating psychological signatures with text analysis. *Psychol. Sci.* **26**, 570–582 (2015).
50. J. W. Pennebaker, L. A. King, Linguistic styles: Language use as an individual difference. *J. Pers. Soc. Psychol.* **77**, 1296–1312 (1999).
51. I. Witten, E. Frank, M. Hall, C. Pal, *Data Mining: Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, ed. 4, 2017).
52. M. L. Jockers, *Macroanalysis: Digital Methods and Literary History* (University of Illinois Press, 2013).

Acknowledgments: We thank R. J. Booth for assistance in the design the Arc of Narrative website. We are also grateful to J. Malin and all of the research assistants and colleagues who have given their time and valuable insights into this project over the years. **Funding:** Preparation of this manuscript was aided by grants from the NIH (5R01GM112697-02 and 1R01MH117172), the John Templeton Foundation (#48503 and #62256), the High-Value Detainee Interrogation Group (15F06718R0006603), and the NSF (IIS-1344257 and SES1758835). This research was also supported by an appointment to the Intelligence Community Postdoctoral Research Fellowship Program at The University of Texas at Austin, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Office of the Director of National Intelligence. **Author contributions:** R.L.B., K.G.B., and J.W.P. were all involved in conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, writing (original draft, reviews, revisions, and editing), and visualization, to varying degrees each. **Competing interests:** J.W.P. is the owner of the LIWC2015 software; all profits from LIWC2015 sales are donated to the University of Texas at Austin. All other authors declare no competing interests. **Data and materials availability:** Supporting data, materials, etc., can be found on our OSF repository, located at <https://osf.io/q2a7m/>. The corresponding author may be contacted for any additional queries.

Submitted 14 November 2019

Accepted 26 June 2020

Published 7 August 2020

10.1126/sciadv.aba2196

Citation: R. L. Boyd, K. G. Blackburn, J. W. Pennebaker, The narrative arc: Revealing core narrative structures through text analysis. *Sci. Adv.* **6**, eaba2196 (2020).

The narrative arc: Revealing core narrative structures through text analysis

Ryan L. Boyd, Kate G. Blackburn and James W. Pennebaker

Sci Adv **6** (32), eaba2196.
DOI: 10.1126/sciadv.aba2196

ARTICLE TOOLS

<http://advances.sciencemag.org/content/6/32/eaba2196>

SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2020/08/03/6.32.eaba2196.DC1>

REFERENCES

This article cites 22 articles, 3 of which you can access for free
<http://advances.sciencemag.org/content/6/32/eaba2196#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).