Historical language records reveal a surge of cognitive distortions in recent decades

Bollen, Johan ; ten Thij, Marijn; Breithaupt, Fritz ; Barron, Alexander T. J. ; Rutter, Lauren A. ; Lorenzo-Luaces, Lorenzo ; Scheffer, Marten

DOI
10.1073/pnas.2102061118

Publication date
2021

Document Version
Final published version

Published in
Proceedings of the National Academy of Sciences of the United States of America

Citation (APA)
Bollen, J., ten Thij, M., Breithaupt, F., Barron, A. T. J., Rutter, L. A., Lorenzo-Luaces, L., & Scheffer, M. (2021). Historical language records reveal a surge of cognitive distortions in recent decades. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(30), 1-7. Article e2102061118. https://doi.org/10.1073/pnas.2102061118

Historical language records reveal a surge of cognitive distortions in recent decades

Johan Bollen[a,1], Marijn ten Thij[a,b], Fritz Breithaupt[c], Alexander T. J. Barron[a], Lauren A. Rutter[d], Lorenzo Lorenzo-Luaces[d], and Marten Scheffer[e]

[a]Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47408; [b]Delft Institute of Applied Mathematics, Delft University of Technology, 2628 CD Delft, The Netherlands; [c]Department of Germanic Studies, Indiana University, Bloomington, IN 47405; [d]Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405; and [e]Department of Environmental Sciences, Wageningen University, 6708 PB Wageningen, The Netherlands

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved June 15, 2021 (received for review February 1, 2021)

Individuals with depression are prone to maladaptive patterns of thinking, known as cognitive distortions, whereby they think about themselves, the world, and the future in overly negative and inaccurate ways. These distortions are associated with marked changes in an individual's mood, behavior, and language. We hypothesize that societies can undergo similar changes in their collective psychology that are reflected in historical records of language use. Here, we investigate the prevalence of textual markers of cognitive distortions in over 14 million books for the past 125 y and observe a surge of their prevalence since the 1980s, to levels exceeding those of the Great Depression and both World Wars. This pattern does not seem to be driven by changes in word meaning, publishing and writing standards, or the Google Books sample. Our results suggest a recent societal shift toward language associated with cognitive distortions and internalizing disorders.

cognitive distortions | internalizing disorders | historical language analysis

Depression is a leading contributor to the burden of disability worldwide (1, 2). Some evidence indicates that disability attributed to depression has been rising over the past decades, particularly among youth (3–5). Can societies collectively become more or less depressed over time, as their populations are exposed to stressors such as war, political upheaval, economic collapse, food insecurity, inequality, and disease (6, 7)? This question is difficult to answer for long time scales because formal diagnostic criteria were introduced only 40 y ago and these criteria have undergone changes over time (8).

Depression is associated with distinct and recognizable maladaptive thinking patterns, referred to as cognitive distortions, wherein individuals think about themselves, the future, and the world in inaccurate and overly negative ways (9–12). For example, a cognitive distortion seen in depression occurs when individuals label themselves in negative, absolutist terms (e.g., "I am a loser"). They may talk about future events in dichotomous, extreme terms (e.g., "My meeting will be a complete disaster") or make unfounded assumptions about someone else's state of mind (e.g., "Everybody will think that I am a failure"). Typologies of cognitive distortions generally differentiate between a number of partially overlapping types, such as "catastrophizing," "dichotomous reasoning," "disqualifying the positive," "emotional reasoning," "fortune telling," "labeling and mislabeling," "magnification and minimization," "mental filtering," "mindreading," "overgeneralizing," "personalizing," and "should statements."

The theory underlying cognitive-behavioral therapy (CBT), the gold standard for the treatment of depression and other internalizing disorders (13), holds that cognitive distortions are associated with internalizing disorders; they reflect negative affectivity and avoidant behavioral patterns in the context of environmental stress (14, 15). Language is closely intertwined with this dynamic. In fact, recent research shows that individuals with internalizing disorders express significantly higher levels of cognitive distortions in their language (16, 17) to the point that their prevalence may be used as an index of vulnerability for depression (18, 19).

Here, we leverage the connection between depression and language to investigate whether societies as a whole, similar to individuals with depression, can undergo changes in their collective language that are associated with cognitive distortions. We analyze the prevalence of a large set of markers of cognitive distortions over the past 125 y in a collection of more than 14 million books (Google Books) published in English, Spanish, and German. Specifically, we are examining the longitudinal prevalence of hundreds of short sequences of one to five words (n-grams), labeled cognitive distortion schemata (CDS), that were designed by a team of CBT experts, computational linguists, and bilingual native speakers and externally validated by a panel of CBT experts, to capture the expression of 12 types of cognitive distortions (9). The CDS n-grams were designed as short, unambiguous, and stand-alone statements that expressed the core of a particular cognitive distortion type, using highly frequent terms (Fig. 1 and *SI Appendix*, Tables S1–S3). For example, the 3-gram "I am a" captures a labeling and mislabeling distortion, regardless of its context or the precise labeling involved ("lady," "honorable person," "loser," etc.). These same n-grams were in earlier research shown to be significantly more

Significance

Can entire societies become more or less depressed over time? Here, we look for the historical traces of cognitive distortions, thinking patterns that are strongly associated with internalizing disorders such as depression and anxiety, in millions of books published over the course of the last two centuries in English, Spanish, and German. We find a pronounced "hockey stick" pattern: Over the past two decades the textual analogs of cognitive distortions surged well above historical levels, including those of World War I and II, after declining or stabilizing for most of the 20th century. Our results point to the possibility that recent socioeconomic changes, new technology, and social media are associated with a surge of cognitive distortions.

Author contributions: J.B., F.B., A.T.J.B., L.A.R., L.L-L., and M.S. designed research; J.B., M.tT., F.B., A.T.J.B., L.A.R., and L.L-L. performed research; J.B. and M.tT. analyzed data; and J.B., M.tT., F.B., A.T.J.B., L.A.R., L.L-L., and M.S. wrote the paper.

Competing interest statement: L.L-L. received an honorarium for consulting from Happify, Inc. in September 2020. Happify had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: jbollen@indiana.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2102061118/-/DCSupplemental.

Published July 23, 2021.

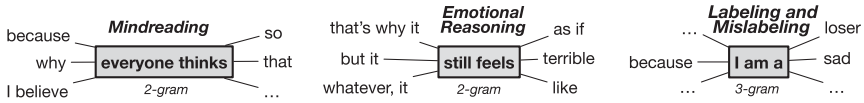

Fig. 1. Examples of CDS n-grams shown inside gray boxes, surrounded by plausible context words that may vary without affecting whether the n-gram marks the expression of a cognitive distortion of the given type (e.g., mindreading, emotional reasoning, or labeling and mislabeling). CDS were designed by a team of CBT experts, linguists, and native language speakers to capture the expression of a particular cognitive distortion type, regardless of its specific lexical context. For English (US), Spanish, and German the team of experts defined respectively 241, 435, and 296 n-grams to mark 12 commonly distinguished types of cognitive distortions. Note that our prevalence measurements count only the CDS n-gram occurrence regardless of context ("everyone thinks," "still feels," and "I am a"). A complete list of all CDS n-grams by distortion type is provided in *SI Appendix*, Tables S1–S3.

prevalent in the language of individuals with depression vs. a random sample (17).

To account for changes in publication volume, for each CDS n-gram we define its prevalence in a given year as the number of times it occurred that year in the Google Books data divided by the total volume published (estimated from end-of-sentence punctuation numbers). All resulting time series are converted to z scores, to provide the same scale of comparison between different CDS n-grams, and compared to a null model of randomly chosen n-grams for the same years and set of books (*Materials and Methods*).

Results

We perform this analysis for three unique geographic and linguistic spheres: 1) the United States of America (US English), 2) the German-speaking countries (German), and 3) all Spanish-speaking countries (Spanish). English (US), Spanish, and German were chosen as the focal points of our analysis because they share a common alphabet, a common history and vary in terms of whether they are spoken as a first language either in a particular geographic region (US books only and German-speaking countries) or across several continents (Spanish) as a control. We limit our analysis to the range 1855 to 2019 since it provides 125 y of persistently high publishing volume for all three languages and few grammatical, orthographic, or spelling changes that would affect our analysis. Although books published in a specific language in a particular geographic area are not necessarily a representative reflection of society as a whole, persistent language trends over decades and centuries, observed from tens of millions of books, have in previous research been shown to signal cultural, linguistic, and psychological changes (18, 20–28).

Trends for English (US), Spanish, and German. We first examine the history of the median prevalence (z scores) of the entire set of English cognitive distortion schemata ($n = 241$) in English (US) books (N = 9,018,119, United States only), from 1855 to 2019 (Fig. 2*A*). Since these data pertain only to books published in the United States, we mark notable events in US history or notable changes in the time series: the end of the century in 1899; the start of World War I; the financial collapse of 1929; the start of World War II; a peak of CDS prevalence in 1968; and distinct trend changes in 1978, 1999, and 2007.

The overall trend of CDS prevalence for most of the 20th century pointed distinctly downward toward a historic minimum in 1978, with only a few noticeable peaks, one surrounding the turn of the century in 1899 (possibly related to the Spanish–American war), a slight peak from 1940 to 1945 (around the time of World War II), and a sharp peak in 1968 (possibly related to social and political unrest). From 1978, we observe an accelerating increase in CDS prevalence. This acceleration seems to be separated into three periods: an accelerating increase from 1978 to 1999 (where CDS prevalence first exceeds levels observed in the 1910s), an even more rapid increase after 1999 to roughly 2007, followed by an acceleration after 2007, and a possible stabilization in 2010. The so-called "bursting of the dot.com bubble" seems to coincide with an acceleration of the increase of CDS prevalence after 1999 whereas the acceleration since 2007 seems to coincide with the widespread uptake of social media and the start of the Great Recession. Present CDS prevalence levels exceed those observed since the 1900s by almost two standard deviations (excepting the 1899 peak).

We include Spanish in our analysis as a control vs. English (US) and German since it is not confined to a particular geographic region; i.e., these data encompass all books published in Spanish, which includes Spain (Europe) and most of Latin America ($N = 1{,}658{,}438$ books). The prevalence of Spanish CDS markers ($N = 435$ n-grams) remains quite stable throughout the 20th century, with a moderate increase around the start of World War 1, a very short moderate spike in 1929, and an upward

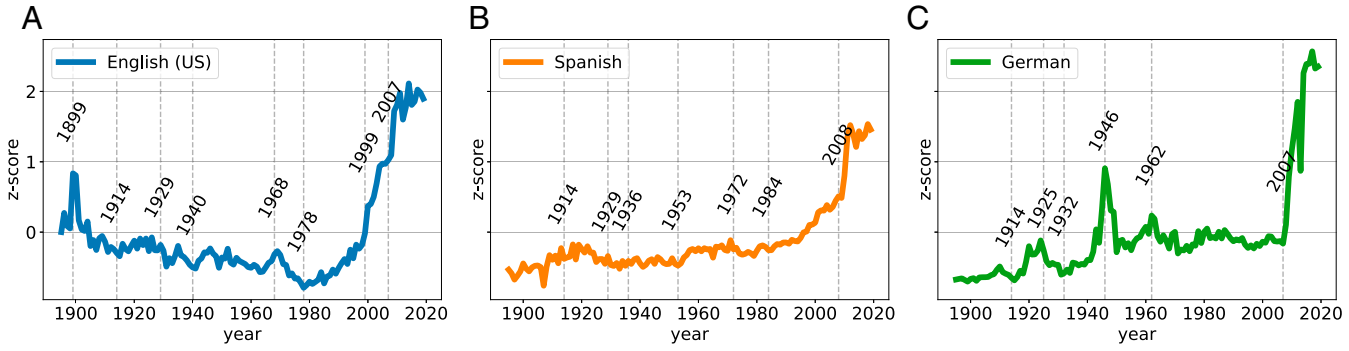

Fig. 2. (*A–C*) Median z scores of time series of CDS n-gram prevalence from 1855 to 2020 (125 y) in US English (*A*), Spanish (*B*), and German (*C*) with year markers added for major historical events. All time series reveal stable or declining levels for most of the 20th century followed by a sharp surge of cognitive distortions in the past three decades. US English shows declining levels from 1899 to 1978, with minor peaks around 1914 and 1940 (World War I and World War II) and notably 1968. This decline is followed by a surge of CDS prevalence starting in 1978 that continues to 2019. For Spanish we find stable levels from 1895 to the early 1980s at which point a trend occurs toward higher CDS prevalence levels above any of those previously observed. German shows stable CDS prevalence levels, with the exception of strong peaks around and after World War I and World War II, until 2007 at which point a sudden surge occurs.

departure from a 30-y downward trend in 1953 after which levels seem to level off until 1984 (Fig. 2*B*). Starting in 1984, however, we observe the same hockey-stick pattern as we saw for English (US): a sharp acceleration of an upward trend starting in 1984 leading to present CDS prevalence levels that exceed the historical baseline by more than one standard deviation. This trend seems to accelerate in 2008.

The pattern of prevalence for German (Fig. 2*C*) books (N = 3,843,962) provides face validity for the ability of our CDS markers ($N = 296$ n-grams) to capture significant moments of stress in a population, since they match major historical and geopolitical events that are specific to German-speaking countries (predominantly Germany and Austria). Contrary to English and Spanish CDS, prevalence levels start relatively low around the 1900s, but sharply increase since the start of World War I, reaching a peak in 1920 and 1923, coinciding with the aftermath of World War I in Germany and a devastating recession in 1923. Throughout the existence of the Weimar Republic, we observe decreasing levels of CDS markers.

However, this trend is interrupted in 1932 at which point cognitive distortion levels increase sharply. This period includes major social upheaval, economic struggles, the end of the Weimar Republic, the emergence of the Nazi regime, and the start of World War II. CDS prevalence levels increase rapidly during World War II, reaching their peak in 1946, the year after Germany was defeated. CDS prevalence declines sharply afterward and reaches a stable plateau throughout the 1950s to 2007, with only a minor peak in 1962 and no indications of accelerating CDS prevalence levels during the 1970s or 1980s as we observe for English (US) and Spanish. Notably, in 2007, at the start of the worldwide Great Recession we see a nearly immediate increase of CDS prevalence levels to nearly two standard deviations above the historical mean.

Null Model of Randomly Chosen N-Grams. We compare the pattern of change for all three languages in terms of the 95% confidence intervals of CDS prevalence (*Materials and Methods*, *Bootstrapping*) for English (US), Spanish, and German against a null model of 10,000 sets of 241 randomly chosen n-grams (Fig. 3). These sets of random n-grams were sampled from all n-grams in the respective English, Spanish, and German Google n-gram corpus such that they have the same number of 1- to 5-grams as the respective CDS set and the same bias toward recently published books due to increased publication volume over time (*Materials and Methods*, *Null Model*).

We provide annotations that mark significant historical events in the graph that have affected the three populations such as the financial crisis of 1929 ("Wall St. crash"), the two World Wars, and the great recession starting in 2007. CDS prevalence levels for English (US), Spanish, and German significantly exceed those of this null model in recent decades, but in the case of Germany also during World War I and World War II. Note that English (US) levels fall below those of the null model from the 1920s to the 1990s (*SI Appendix*, Fig. S6).

Trends for Distinct Cognitive Distortion Types. We plot the time series of yearly mean CDS prevalence separated by 12 commonly recognized cognitive distortion types (14) for English (US), Spanish, and German (Fig. 4). For all three languages and across most cognitive distortion types we see the characteristic hockey-stick signature of stable or declining CDS prevalence levels followed by a surge above historical levels during the period 1980 to 2010 to levels well above the historical mean. The one exception is should statements, which, due to their grammatical structure, may be difficult to translate to n-grams uniquely associated with the specific cognitive distortion.

For English we frequently see a "tilted hockey stick" pattern where certain types of CDS n-grams declined over the 20th century followed by a rapid surge of prevalence since 1978. This is the case for fortune telling, overgeneralizing, magnification and minimization, mindreading, and labeling and mislabeling and perhaps most pronounced for dichotomous reasoning, suggesting that these distortion types are likely responsible for the decline of CDS prevalence observed throughout the 20th century (Fig. 2*A*). We also find slight peaks for catastrophizing, emotional reasoning, and mindreading at the time of the US involvement in World War II. For German, however, we see peaks surrounding World War I and World War II for dichotomous reasoning, fortune telling, labeling and mislabeling, mental filtering, mindreading, overgeneralizing, personalizing, and should statements, possibly indicating the widespread effects of the two World Wars on German language use.

Robustness and Limitations. As our observations could be caused by a number of effects and biases, we conducted a number of mitigation controls and sensitivity analyses to test alternative explanations for the observed patterns.

Language effects. We caution that changes in meaning or semantic shift of the CDS n-grams may potentially bias our results. As a rhetorical example, the 1-gram "ever" could acquire a different meaning or usage over time and hence lose its meaning as a marker of a cognitive distortion of the dichotomizing

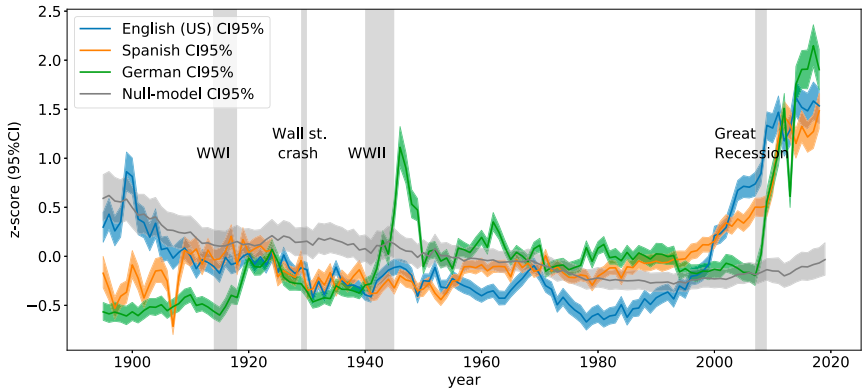

Fig. 3. CDS prevalence for English, Spanish, and German superimposed with a null-model estimate of random n-gram prevalence. Colored bands indicate 95% confidence intervals of yearly z-score values estimated with 10,000-fold bootstrap of the set of individual CDS time series. Gray band indicates 95% confidence interval of a null model of 10,000 sets of 241 randomly chosen n-grams with the same length distribution as the English (US) CDS set (*Materials and Methods*, *Null Model*).

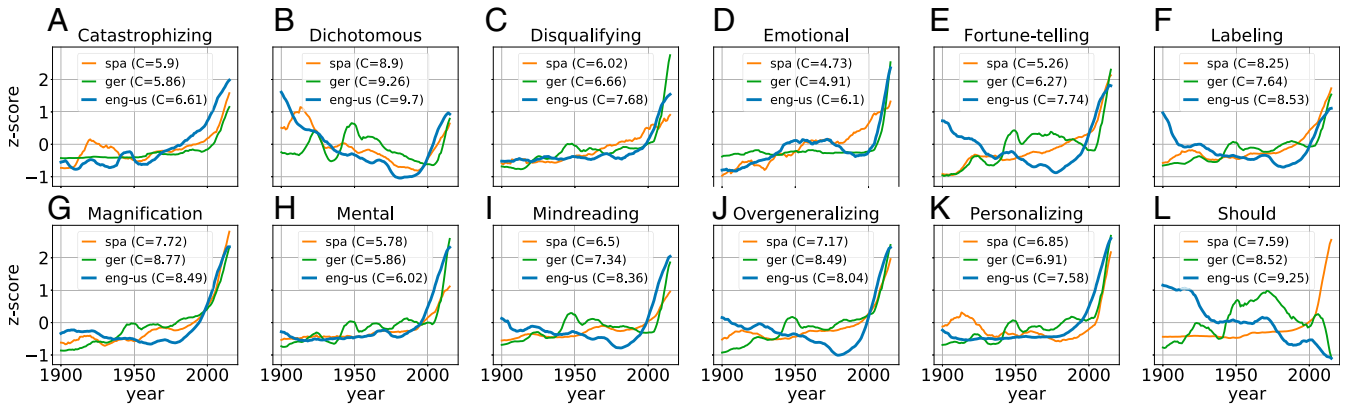

Fig. 4. (*A–L*) CDS n-gram prevalence from 1855 to 2019 (median z score smoothed by 10-y rolling mean), for English, Spanish, and German, grouped by cognitive distortion type, namely (*A*) catastrophizing, (*B*) dichotomous reasoning, (*C*) disqualifying the positive, (*D*) emotional reasoning, (*E*) fortune telling, (*F*) labeling and mislabeling, (*G*) magnification and minimization, (*H*) mental filtering, (*I*) mindreading, (*J*) overgeneralizing, (*K*) personalizing, and (*L*) should statements. Nearly all time series reveal a universal hockey-stick pattern of recently surging CDS n-gram prevalence levels across cognitive distortion types. The value *C* indicates the log (base 10) of the total frequency of CDS n-grams in the specific cognitive distortion category as an indication of the order of magnitude of its contribution to our observations.

type. We perform several controls to account for changes in language over time. First, the CDS n-grams consist predominantly of words that have been among the most frequent since 1895 [mean word percentile among all 1-grams $M(P_r) = 99.885, \mathrm{SD} = 0.346$; *SI Appendix*, Fig. S1]. The CDS n-grams have equally been among the most frequent since 1895 [mean CDS n-gram percentile among all 2- to 5-grams $M(P_r) = 0.946, \mathrm{SD} = 0.010$; *SI Appendix*, Fig. S2]. Hamilton, Leskovec, and Jurafsky (25) quantify semantic shifts over historical time using word embeddings, showing that frequent words experience the lowest rate of change, scaling with an inverse power law of word frequency. Hence the rate of semantic shift of the words in our CDS n-grams and the CDS n-grams themselves could be among the lowest as well. Second, a trend toward shorter sentences (29) may provide alternative explanations of our observations, but although sentence length did decrease from 1890 to the 1920s in English, it has remained stable since (30). Furthermore, our analysis accounts for changes in sentence length by normalizing n-gram prevalence by the frequency of end-of-sentence punctuation for that year (*Materials and Methods*, *Time Series: Prevalence and Normalization*). Finally, we previously showed that the prevalence of the CDS n-grams in the language of individuals with depression is not affected by the emotional valence of the n-grams or the presence of personal pronouns (17); hence, a language trend toward more emotional language or use of personal pronouns is not likely to affect our results.

Sampling effects. There are several issues that could arise from our Google Books sample. First, Pechenick et al. (31) show indications of a possible increase in technical writing and nonfiction in the Google Books sample over the past decades. Since our CDS n-grams contain personal pronouns, common verbs, and adjectives that may refer to personal matters, if the amount of technical writing and nonfiction increased, one could hypothesize this could explain a decrease in CDS prevalence. However, we observe the opposite, a significant increase of CDS prevalence.

Second, the choice of CDS n-grams could lead to a "recency bias" in our results, explaining their rise in prevalence in recent decades. We control for this effect with a null model that samples random n-grams more frequently from recent books, due to rapidly increasing publication volume since 1895, thereby inducing a bias toward more recent language. We observe increases of CDS n-gram prevalence well above levels predicted by this null model (Fig. 3). Hence, a recency bias alone may not likely explain the observed surge in CDS prevalence in recent decades relative to this null model.

Finally, all n-grams in the English (US), Spanish, and German CDS sets occurred in every year from 1895 to 2019, indicating they were in continuous use throughout this period. They were highly frequent from 1895 to 2019, in fact on average more frequent than 94.6% ($\mathrm{SD} = 0.0103$) of all n-grams in the Google Books data (*SI Appendix*, Figs. S1 and S2). We furthermore bootstrap our prevalence estimates to gauge the sensitivity of our findings to random changes in the set of CDS n-grams over time (*Materials and Methods*, *Bootstrapping*). The narrow 95% CI bands (Fig. 3) throughout the period under consideration indicate the stability of our observations over time.

CDS limitations. We caution that although the Google Books data have been widely used to assess cultural and linguistic shifts, and they are one of the largest records of historical literature, it remains uncertain whether CDS prevalence truly reflects changes in societal language and societal wellbeing. Many books included in the Google Books sample were published at times or locations marked by reduced freedom of expression, widespread propaganda, social stigma, and cultural as well as socioeconomic inequities that may reduce access to the literature, potentially reducing its ability to reflect societal changes. Although CDS n-gram prevalence was shown to be higher in individuals with depression (17) and our composition of CDS n-grams closely follows the framework of cognitive distortions established by Beck (9), they do not constitute an individual diagnostic criterion with respect to authors, readers, and the general public. It is also not clear whether the mental health status of authors provides a true reflection of societal changes nor whether cultural changes may have taken place that altered the association between mental health, cognitive distortions, and their expression in language.

Discussion

While the differences between the languages are interesting, perhaps the most important point is that the expression of cognitive distortions increases for all three languages in the recent three decades, leading to a distinct hockey-stick pattern indicating a surge of the CDS prevalence levels, which serve as lexical markers of cognitive distortions.

We can only speculate on the possible underlying causes of the observed surge of CDS prevalence for these three languages, since our results do not establish any causal mechanisms. The

strong increases of CDS prevalence in German during World Wars I and II are validating with respect to the ability of our CDS n-grams to signal societal dynamics in times of turmoil and run counter to the hypothesis that our results are caused by a recency bias in our choice of CDS n-grams and the Google Books sample. In fact, the surge of CDS prevalence during and right after World War II may be the product of a detrimental combination of the war experience and National Socialist propaganda. While there was not a separate language of National Socialism (32), the discourse of National Socialism invaded many registers of speech, including everyday language use, and thus normalized the political agenda (33, 34). In particular, the discourse of National Socialism is shaped by a language of identity that emphasizes an us-them divide (35), which relates to several markers of CDS n-grams, e.g., dichotomous reasoning, labeling and mislabeling, mindreading, and fortune telling. The tumultuous period of 1959 to 1962 in Germany cemented the division of East and West when the Berlin Wall was built in 1961. Interestingly, the fall of the Wall in 1989 did not result in higher amounts of psychological distress (36) and also does not register on fluctuations of CDS n-grams prevalence. The German data show a period of stability from 1962 up to the rapid increase after 2007.

Other differences between the dynamics of cognitive distortions in the three language corpora we analyzed might also point to relevant drivers. For instance, in Spanish and English we see a rising trend starting around 1980, whereas in German there is no such rise, only a sharp jump to a higher level in 2007. It is possible that the reunification of Germany in 1990 and the increased integration of the German-speaking countries in the European Union (and the introduction of the Euro currency in 1999) provided resilience to trends recorded in Spanish and English prior to 2007.

It is suggestive that the timing of the US surge in CDS prevalence coincides with the late 1970s when wages stopped tracking increasing work productivity. This trend was associated with rises in income inequality to recent levels not seen since the 1930s (37). This phenomenon has been observed for most developed economies, including Germany, Spain, and Latin America, contemporaneous with the rapid growth of automation and demand for highly skilled labor (38). The great recession of 2007 might have compounded the effects of this decades-long trend that started in the late 1970s. The widespread adoption of communication technologies such as the internet, the World Wide Web, and social media (39–42) may have driven greater societal and political polarization (43, 44) at a global level (45). The language of such polarization may correspond to cognitive distortions (46), in particular us-vs.-them thinking (labeling and mislabeling), dichotomous reasoning, mindreading (47), overgeneralizing, emotional reasoning, and catastrophizing.

We caution that we make no causal claims with respect to the relationship between lexical markers, cognitive distortions, and internalizing disorders, and the above comments therefore constitute speculations that we hope may inspire follow-up research. Regardless of speculation with respect to the underlying cultural, social, or economic drivers, our results indicate historically high levels of the expression of a large set of lexical markers of cognitive distortions in three languages. Given the association between cognitive distortions and internalizing disorders, this points to the possibility that large populations are increasingly stressed by pervasive cultural, economic, and social changes. The rising prevalence of depression and anxiety (3–5) in recent decades seems to align with our observations.

The availability of large-scale historical records of published languages going back centuries may provide a unique opportunity for the quantitative investigation of important cultural and linguistic dynamics ("culturomics") (21), while acknowledging limitations with respect to verifying hypotheses and testing the causal mechanisms that underlie any observations from these data. Future work may contribute to a better understanding of how changes in the collective psychology of societies can be observed over time and how these changes are manifested in their language in response to a variety of cultural and socioeconomic challenges, for example from quantitative indicators of semantic shifts in word meaning (25).

Materials and Methods

Google Books N-Gram Data. We used the third version 2019 release of the Google Books n-gram data, which the Google Books team makes freely available online (https://storage.googleapis.com/books/ngrams/books/datasetsv3.html). The data span from the 16th century to the year 2019, with increasing coverage of later years as publication volume grew rapidly.

Cognitive Distortion N-Grams. A panel of CBT experts engaged in a collaborative design effort to compile a set of 241 English n-grams (sequences of n = 1, 2, 3, 4, and 5 words) that were deemed to indicate the expression of a cognitive distortion of a particular type. These CDS n-grams were designed to consist of simple, frequent, and nontechnical expressions, e.g., "I am a," "he thinks," "will be," etc., designed to be stand-alone expressions of cognitive distortions regardless of their specific context. For example, "I am a [defeated gentlemen]" and "I am a [loser]" both constitute an expression of a labeling and mislabeling captured by the "I am a" 3-gram.

The set of 241 English n-grams was subsequently translated from English into Spanish and German. This translation mainly focused on retaining the cognitive distortion expression of an n-gram in the target language. All translations were collaboratively compared, back translated, and validated by members of our team of CBT experts and native language speakers to ensure consensus. Since CDS n-grams are short expressions of frequent one, two, three, four, or five words, many have nearly literal translations, e.g., "I am a" translates to "Yo soy un" and "yo soy una." We provide the complete lists of all English, Spanish, and German CDS n-grams in *SI Appendix*, Tables S1–S3.

The number of CDS n-grams is higher in Spanish and German than in English since the former need to capture grammatical variations such as conjugations, cases and inflections, and gender. Some n-grams were translated to regular expressions (RE) to capture succinctly all possible lexical and grammatical variations of the same CDS in Spanish and German. For example, the English "I am a" was translated to the REs "Yo soy un(a)," which matches both the male and female gender of the definite article in Spanish, and "ich bin ein(e,em,er,en,es)" to match all possible grammatical variations in German (including misspellings or errors). Note that in the case of German the operative verb or term is frequently at the end of the sentence. For example, the n-gram "I never" was translated to the regular expression "Ich (.+) nie," matching any 3-, 4-, and 5-gram that starts with "Ich" (I) and ends in "nie" (never) such as "Ich habe nie" (I never have), "Ich hatte nie" (I never had), and "Ich war nie" (I never was).

All n-grams and REs were matched in a case-insensitive manner against the Google Books data to capture the widest possible lexical variation across time, including all capitalizations. One given CDS, depending on capitalization and grammatical variations, could match multiple expressions. For example, the German regular expression "Ich bin ein(e,es,em,. . .)" could match any of "ich bin ein," "Ich bin eine," "ich bin eines", etc., including some variations that are rare or even grammatically incorrect (e.g., other letters capitalized). All such matches for each individual n-gram were summed into a single compound time series for the specific CDS such that the most frequent forms have the highest relative weight in subsequent analysis.

Each match retrieved the complete time series of n-gram frequencies for the specific n-gram and matching RE from the earliest to the latest year provided. The time series values correspond to the number of times the n-grams occurred in the Google Books data in a particular year. Note that the earliest books in the Google Books dataset were published in the 15th century, a period marked by low publication volumes and variances in spelling and grammar. As mentioned, our analysis was limited to the period 1895 to 2019 to capture the end of the 19th century, most of the 20th century, and the past two decades, a period of high publishing volume, relatively stable orthographic standards, and relatively low variance of our CDS prevalence data (*SI Appendix*, Fig. S4). Each n-gram in the CDS set of each language occurred in every year from 1895 to 2019, indicating continuous coverage for all individual CDS n-grams.

Time Series: Prevalence and Normalization. The volume of books published has increased significantly over the past two centuries, punctuated by declines at times of economic collapse and war (*SI Appendix*, Fig. S3). The frequency of occurrence of any specific n-gram will therefore fluctuate accordingly, since it is recorded from a changing sample of books published. We therefore determine yearly n-gram prevalence by normalizing the observed yearly frequency of each n-gram by the total yearly volume published. We estimate the latter by summing the yearly frequency of periods, exclamation points, and question marks ("." "!," and "?"), three punctuation symbols that are used in English, Spanish, and German to mark the end of a sentence. Their frequency indicates publication volume as the number of sentences published (*SI Appendix*, Fig. S4), accounting for possible changes in writing style toward shorter or longer sentences. Although periods, exclamation points, and question marks may express different meanings over time, here we record their frequency only to mark the end of sentences. The ratio of periods vs. all end-of-sentence punctuation remained stable from 1895 to 2019 (M = 0.9633, SD = 0.0103), with periods approximately 26 times more frequent than exclamation points and question marks.

More formally we determine our prevalence time series as follows. We define a set C of k n-grams $c_i \in C = \{c_1, c_2, \ldots, c_k\}$, where the number of words in each n-gram is its length $n \in \{1, 2, 3, 4, 5\}$. We denote the yearly time series of the prevalence of any n-gram c_i as the set $X_j(c_i)$, where j refers to a year in the ordered set (1855, 1856, . . . , 2020). Each value $x_j \in \mathbb{N}^+$ of $X_j(c_i)$ represents the n-gram's prevalence, i.e., the total number of occurrences of n-gram c_i in the books published in year j, which we denote $x_j(c_i)$.

Since the raw prevalence $x_j(c_i)$ will fluctuate with the total volume of text published in a given year j, denoted V_j, we normalize the prevalence $x_j(c_i)$ with an estimate of V_j. We use the total number of yearly occurrences of end-of-sentence punctuation to estimate the volume of text published, $N_j = X_j(.) + X_j(!) + X_j(?)$; hence, the volume-normalized time series of n-gram c_i is given by $\bar{X}_j(c_i) = X_j(c_i)/N_j$ for every year j. Note that N_j roughly corresponds to the total number of sentences published, because nearly all sentences are terminated by end-of-sentence punctuation.

However, the magnitude of yearly normalized prevalence values will differ significantly between n-grams of different lengths; e.g., "never" is likely much more prevalent than "they will not believe" for the same volume published, since the former is a 1-gram and the latter a more specific 4-gram. Nevertheless, both may follow a similarly shaped pattern of historical change. Furthermore, the volume of books published increases rapidly over time; hence the variance of the time series of n-gram occurrence may also change over time. To allow comparisons of the patterns of changing prevalence over time between time series of different magnitudes and variance, we subtract the 1895 to 2019 mean from all prevalence time series and divide them by the observed standard deviation over the same period, thus converting all prevalence time series to z scores with respect to their 1895 to 2019 mean $\mu(\bar{X}_j)$ and standard deviation $\sigma(\bar{X}_j)$ as follows: $Z_j(c_i) = (\bar{X}_j(c_i) - \mu(\bar{X}_j(c_i)))/\sigma(\bar{X}_j(c_i))$.

These time series express the fluctuations of the prevalence of all n-grams on a common scale, namely standard deviations from their historical mean, without altering the pattern of their decline or increase of time. Normalized as such we can compare the pattern of changing prevalence for any CDS n-gram time series on a common scale. For example, the individual time series of the "I am a" n-gram that marks a labeling and mislabeling cognitive distortion can have very different magnitudes in Spanish and English, but similar patterns of change over time. This is revealed when we plot them at the same z-score scale (*SI Appendix*, Fig. S5).

Null Model. We define a null model to compare the observed yearly CDS n-gram prevalence fluctuations against. This null model consists of 10,000 sets of k randomly selected n-grams sampled uniformly across the set of n-grams in the Google Books data. To match the continuity of coverage in the CDS n-grams since 1895, each random n-gram was required to occur in at least 100 of 125 y in our analysis period. Each of the resulting 10,000 random sets of n-grams, denoted $\hat{C}_i \in \hat{C} = \{\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_{10,000}\}$, was chosen to replicate the number of n-grams of a given length n as in C (e.g., there are 86 3-grams in C of $k = 241$ total in English [*SI Appendix*, Table S1], so every $\hat{C}_i$ would also have 86 randomly selected 3-grams of 241 total). We retrieve the Google Books time series for each individual random set of CDS $\hat{C}_i$ and normalize them to prevalence values as described above. This results in 10,000 time series that yield a yearly distribution of z scores. We use the 2.5th and 97.5th percentiles of this yearly distribution as a 95% confidence interval showing the diachronic fluctuations of random n-grams in our data, to which we can compare the empirical fluctuations of our selected set of CDS n-grams (Fig. 3 and *SI Appendix*, Fig. S6).

This null model controls for inherent methodological biases in the normalization of our time series and the Google Books data sample, as well as a possible recency bias in the CDS n-grams since the null model's n-grams are randomly chosen from the Google Books data, which have grown rapidly in volume since 1895 (*SI Appendix*, Fig. S3). Therefore, any recent increase in CDS prevalence (e.g., the observed hockey-stick pattern) is compared against a null model that draws n-grams preferentially from recently published books. In addition, the requirement for each null-model n-gram to have at least 100 y of coverage since 1895 also favors more recent n-grams because more data will be available in recent years.

Bootstrapping. Each CDS n-gram corresponds to a yearly z-score time series from 1855 to 2019, yielding a distribution of z scores for each year for the set of CDS n-grams (*SI Appendix*, Fig. S7). To determine the robustness of our results under random variations of our set of CDS n-grams (e.g., by making different CDS n-gram choices) we calculate the 95% confidence intervals for this distribution by a 10,000-fold random resampling with replacement of the respective set of CDS n-grams (i.e., US English, Spanish, or German). Each such random resample results in a new mean time series for the given random set C_r of n-grams, i.e., $\bar{Z}_{1885,2020}(C_r)$. The resulting yearly distribution of z scores indicates how much our yearly results can vary as a result of random changes to our CDS n-gram set and thereby tests the robustness of our time series results under 10,000 random variations of the set of CDS n-grams. All time series are normalized and converted to z scores before resampling; hence we obtain a distribution of yearly z scores from which the relevant percentiles, including the median, are determined.

Data Availability. Previously published data were used for this work (https://storage.googleapis.com/books/ngrams/books/datasetsv3.html).

ACKNOWLEDGMENTS. J.B. is grateful for support from the Urban Mental Health Institute of the University of Amsterdam, Wageningen University and Research, the NSF (NSF Social, Behavioral and Economic Sciences [SBE] 1636636), and the support of the Indiana University Vice Provost for COVID-19 Research. We thank Jonathan Haidt of New York University who provided the inspiration for this investigation by speculating on a recent societal shift toward a style of thinking that may be associated with internalizing disorders.

1. P. E. Greenberg, A. A. Fournier, T. Sisitsky, C. T. Pike, R. C. Kessler, The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J. Clin. Psychiatr.* **76**, 155–162 (2015).
2. World Health Organization, *Depression and Other Common Mental Disorders: Global Health Estimates* (WHO, Geneva, Switzerland, 2017).
3. A. Case, A. Deaton, Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15078–15083 (2015).
4. R. Mojtabai, M. Olfson, B. Han, National trends in the prevalence and treatment of depression in adolescents and young adults. *Pediatrics* **138**, e20161878 (2016).
5. K. M. Keyes, D. Gary, P. M. O. Malley, A. Hamilton, J. Schulenberg, Recent increases in depressive symptoms among US adolescents: Trends from 1991 to 2018. *Soc. Psychiatr. Psychiatr. Epidemiol.* **54**, 987–996 (2019).
6. J. A. Goldstone *et al*., A global model for forecasting political instability. *Am. J. Polit. Sci.* **54**, 190–208 (2010).
7. J. DeVylder, L. Fedina, B. Link, Impact of police violence on mental health: A theoretical framework. *Am. J. Publ. Health* **110**, 1704–1710 (2020).
8. American Psychological Association, *Diagnostic and Statistical Manual of Mental Disorders* (Am. Psychiatric Assoc., ed. 5, 2013).
9. A. T. Beck, Thinking and depression: I. Idiosyncratic content and cognitive distortions. *Arch. Gen. Psychiatr.* **9**, 324–333 (1963).
10. A. T. Beck, Thinking and depression: II. Theory and therapy. *Arch. Gen. Psychiatr.* **10**, 561–571 (1964).
11. D. Burns, *The Feeling Good Handbook* (Harper-Collins Publishers, 1989).
12. J. S. Beck, A. T. Beck, *Cognitive Therapy: Basics and Beyond* (Guilford Press, New York, NY, 1995).
13. L. Lorenzo-Luaces, The evidence for cognitive behavioral therapy. *Jama* **319**, 831–832 (2018).
14. A. T. Beck, E. A. Haigh, Advances in cognitive theory and therapy: The generic cognitive model. *Annu. Rev. Clin. Psychol.* **10**, 1–24 (2014).
15. D. A. Clark, A. T. Beck, Cognitive theory and therapy of anxiety and depression: Convergence with neurobiological findings. *Trends Cognit. Sci.* **14**, 418–424 (2010).
16. W. Bucci, N. Freedman, The language of depression. *Bull. Menninger Clin.* **45**, 334 (1981).

17. K. C. Bathina, M. Thij, L. Lorenzo-Luaces, L. A. Rutter, J. Bollen, Depressed individuals express more distorted thinking on social media. arXiv:2002.02800 (7 February 2020).
18. M. Al-Mosaiwi, T. Johnstone, In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clin. Psychol. Sci.* **6**, 529–542 (2018).
19. J. C. Eichstaedt *et al.*, Facebook language predicts depression in medical records. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11203–11208 (2018).
20. G. A. Miller, *The Science of Words* (Scientific American Library Series, W. H. Freeman & Co., 1996).
21. J. B. Michel *et al.*, Quantitative analysis of culture using millions of digitized books. *Science* **331**, 176–182 (2011).
22. P. M. Greenfield, The changing psychology of culture from 1800 through 2000. *Psychol. Sci.* **24**, 1722–1731 (2013).
23. M. Davies, Making Google Books n-grams useful for a wide range of research on language change. *Int. J. Corpus Linguist.* **19**, 401–416 (2014).
24. G. Coppersmith, M. Dredze, C. Harman, "Quantifying mental health signals in Twitter" in *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, CLPsych.*, P. Resnik, R. Resnik, M. Mitchell, Eds. (Association for Computational Linguistics [ACL], Stroudsburg, PA, 2014), pp. 51–60.
25. W. L. Hamilton, J. Leskovec, D. Jurafsky, "Diachronic word embeddings reveal statistical laws of semantic change" in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, A. van den Bosch, K. Erk, N. A. Smith, Eds. (Association for Computational Linguistics, Berlin, Germany, 2016), pp. 1489–1501.
26. R. Amato, L. Lacasa, A. Díaz-Guilera, A. Baronchelli, The dynamics of norm change in the cultural evolution of language. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8260–8265 (2018).
27. P. Lorenz-Spreen, B. M. Mønsted, P. Hövel, S. Lehmann, Accelerating dynamics of collective attention. *Nat. Commun.* **10**, 1759 (2019).
28. T. T. Hills, E. Proto, D. Sgroi, C. I. Seresinhe, Historical analysis of national subjective wellbeing using millions of digitized books. *Nat. Human Behav.* **3**, 1271–1275 (2019).
29. R. Iliev, J. Hoover, M. Dehghani, R. Axelrod, Linguistic positivity in historical texts reflects dynamic environmental and psychological factors. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7871–E7879 (2016).
30. K. Rudnicka, "Variation of sentence length across time and genre" in *Studies in Corpus Linguistics* ,R. J. Whitt, Ed. (John Benjamins Publishing Company), vol. 85, pp. 220–240 (2018).
31. E. A. Pechenick, C. M. Danforth, P. S. Dodds, Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PloS One* **10**, 1–24 (2015).
32. P. Von Polenz, *Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart* (Walter de Gruyter, 1999), vol. 3.
33. U. Maas, *"Als der Geist der Gemeinschaft eine Sprache fand": Sprache im Nationalsozialismus. Versuch einer Historischen Argumentationsanalyse* (Springer-Verlag, 2013).
34. V. Klemperer, *The Language of the Third Reich: Lti - Lingua Tertii Imperii: A Philologist's Notebook* (Bloomsbury Academic, 2013).
35. G. Horan, Er zog sich die neue Sprache des Dritten Reiches ueber wie ein Kleidungsstueck: Communities of practice and performativity in national socialist discourse. *Linguist. Online* **30**, 57–80 (2007).
36. M. Achberger, M. Linden, O. Benkert, Psychological distress and psychiatric disorders in primary health care patients in East and West Germany 1 year after the fall of the Berlin Wall. *Soc. Psychiatr. Psychiatr. Epidemiol.* **34**, 195–201 (1999).
37. Economic Policy Institute, *The State of Working America* (Cornell University Press, (2012).
38. UN Department of Economic and Social Affairs, *World Social Report 2020: Inequality in a Rapidly Changing World* (United Nations, 2020).
39. Y. Kelly, A. Zilanawala, C. Booker, A. Sacker, Social media use and adolescent mental health: Findings from the UK millennium cohort study. *EClinicalScienc*e **6**, 59–68 (2018).
40. B. Keles, N. McCrae, A. Grealish, A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents. *Int. J. Adolesc. Youth* **25**, 79–93 (2020).
41. M. G. Hunt, R. Marx, C. Lipson, J. Young, No more fomo: Limiting social media decreases loneliness and depression. *J. Soc. Clin. Psychol.* **37**, 751–768 (2018).
42. J. M. Twenge, J. Haidt, T. E. Joiner, W. K. Campbell, Underestimating digital media harm. *Nat. Human Behav.* **4**, 346–348 (2020).
43. C. A. Bail *et al.*, Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9216–9221 (2018).
44. N. Rodriguez, J. Bollen, Y. Y. Ahn, Collective dynamics of belief evolution under cognitive coherence and social conformity. *PloS One* **11**, 1–15 (2016).
45. T. Carothers, A. O. Donohue, *Democracies Divided: The Global Challenge of Political Polarization* (Brookings Institution Press, 2019).
46. G. Lukianoff, J. Haidt, *The Coddling of the American* (Penguin Books, 2015).
47. K. Barasz, T. Kim, I. Evangelidis, I know why you voted for Trump: (Over)inferring motives based on choice. *Cognition* **188**, 85–97 (2019).